A comparative analysis of six audit systems for mental health nursing

RUTH BALOGH1, HELEN QUINN2, ANNE SIMPSON1 AND SENG A BOND1

1Centre for Health Services Research, University of Newcastle-upon-Tyne, 21 Claremont Place, Newcastle-upon-Tyne NE2 4AA and
2School of Health Sciences, University of Sunderland, Sunderland SR2 3SD, UK

Abstract

Purpose. To devise an analytical framework to help identify strengths and weaknesses in the audit process as specified by existing psychiatric nursing audit systems, in order to analyse current audit practice and identify improvements for incorporation in the Newcastle Clinical Audit Toolkit for Mental Health [1].

Data sources. Published material relating to the following six systems: the Central Nottinghamshire Psychiatric Nursing Audit; Psychiatric Nursing Monitor; Standards of Care and Practice; Achievable Standards of Care; Quartz; and Quest.

Data extraction. Comparison of the six systems according to an analytical framework derived from detailed empirical study (structures, processes and outcomes) of one of them in use and the educational evaluation literature. Examination of the extent to which guidance is provided for operating the systems and for wider process-related aspects of audit.

Results of data synthesis. Five of the systems failed to specify some important elements of the audit process. Conceptually, the six systems can be divided into two main types: ‘instrument-like’ systems designed along psychometric lines and which emphasize the distance between the subjects of audit and the operators of the systems, and ‘tool-like’ systems which exploit opportunities for care setting staff to engage in the audit process. A third type of system is the locally-developed system which is offered to a wider audience but which does not make the same level of claim to universal applicability.

Conclusion. The analytical framework allows different approaches to audit to be compared and contrasted not only according to the techniques used, but also according to process issues. The analysis of six systems revealed a variety of different techniques and procedures which can facilitate, in a methodologically rigorous manner, practitioner and other stakeholder involvement in audit processes.

Keywords: audit process, clinical audit, evaluation, mental health, psychiatric nursing, quality improvement nursing

Quality improvement activities in the UK National Health Service (NHS) have developed over the last decade in response to a growing number of policy requirements which were given a major impetus by the NHS reforms. They have done so out of necessity, and in the absence of a coherent body of literature on theoretical issues to underpin their development as evaluative activities. Reviewing those attempts which have been made to analyse the audit process, Walsh and Coles argue that ‘almost without exception, the criteria for good audit are based on anecdotal evidence at best’ [2].

Questions about the features of an audit or quality assurance (QA) system and how well such systems work have been addressed in the UK mainly by cursory reference to management wisdom rather than to a more rigorous set of ideas.

In nursing, audit has become a well-established element of professional practice, yet the sheer size of the nursing workforce in the UK has meant that many nurses have remained largely untouched by the expansion in audit activities over the past 5 years [3]. One feature of nursing tradition, not only in the UK but internationally, is the use of systematic approaches to audit in which sets of pre-determined topics outlining a comprehensive description of the care process are provided. The Rush-Medicus [4] is an example of such a system which is widely used in the USA and elsewhere.

There are two main advantages to such approaches. Firstly, they attempt to convey a comprehensive portrayal of nursing care. This means that where quality improvements require the redistribution of nursing skills, a balance must be struck so that improvement in one area of practice does not jeopardize other areas of practice where the nursing contribution may be correspondingly reduced. Indeed, some systems incorporate or recommend a workload analysis to facilitate this balance [5]. The provision of global measures also provides the opportunity for care settings to be compared with themselves over time. Secondly, these approaches provide concrete sets of audit topics, thus avoiding the often

Address correspondence to Dr R. Balogh. Tel: (+44) 191 222 7045. Fax: (+44) 191 222 6043.
time-consuming business of identifying topics, developing explicit standards for practice and developing tools for auditing them.

The development of nursing audit as routine clinical practice throughout the UK service still requires a substantial programme of audit activity, and systematic approaches offer a cost-effective method for tackling this deficit. However, these approaches have also attracted criticisms, chiefly relating to the well-documented problems of implementing changes in practice when practitioners have little involvement in the audit process. For example, Harvey’s study shows how pre-determined approaches can lend themselves to implementation in a top-down manner with minimal staff involvement [6]. A further difficulty with these approaches concerns the fixed nature of the audit topics, and hence the frustration of further development within the system once standards have been achieved and maintained.

In seeking to develop a new system for mental health nursing which could be widely applicable both as an introduction to uni-professional audit and as a vehicle for multi-disciplinary audit, we sought to examine the audit process as specified in those systems which were being used in mental health settings in the UK in the early 1990s. Our aim was to clarify differences in approaches to audit and to identify opportunities for staff to be involved in audit projects, thus providing further opportunities for them to learn about the nature of clinical audit and its relationship with clinical practice, for developing their skills in the practice of clinical audit and for implementing appropriate changes to their clinical practice. Our analysis needed to focus on processes and we therefore turned to the literature on educational evaluation where the role of processes in evaluation and evaluative reasoning has been extensively discussed [7,8].

The systems analysed

The systems which were analysed were those in current use in the UK which employed pre-determined sets of audit issues and which aspired to applicability beyond the care settings in which they were developed. Our definition of audit was based on guidance from the English Department of Health which describes clinical audit as the ‘systematic and critical analysis of normal practice against defined purposes and standards, with the introduction of planned changes in response to that analysis’ [9]. It should be noted, however, that in the early stages of clinical audit and QA in the UK the cyclic nature of these activities tended to be used to define them to the exclusion of other activities. For example, the Institute of Clinical Science defined QA as ‘the process in which achievable and desirable levels of quality are described, the extent to which these levels are achieved and the action to enable them to be reached is taken’ [10]. In fact, although QA is an activity of much wider scope, employing a wider variety of tools and mechanisms, in practice there was considerable overlap between the two activities. Therefore, we did not apply a strict definition of clinical audit in identifying the systems for analysis, since this would have excluded systems with similar scope and remit but operating as QA or ‘review’ systems. Instead we examined those systems which had a clear clinical focus and were being used in mental health settings.

We identified six different published systems widely in use in UK psychiatric settings. Four of these had a specific nursing focus, and described themselves as ‘audits’: the Central Nottingham Psychiatric Nursing Audit (CNPNA) [11], Standards of Care and Practice Audit (SCAPA) [12], Quality Evaluation of Standards (QUEST) [13], and Psychiatric Nursing Monitoring Monitor (PNM) [14]. The other two were multi-disciplinary: QUARTZ [15] and Achievable Standards of Care [16]. QUARTZ described itself as a ‘QA system’, while Achievable Standards of Care was described as a ‘review’.

The analytic framework

The study of the use of the CNPNA over a 3-year period in one hospital, and its introduction in another (both in the north of England) provided empirical evidence about the impact of the system on clinical practice [17]. One of the principal findings of this study was that although audits activated nursing staff to implement changes in their practice, the data collected in the course of the audits were often only weakly linked to the recommendations for change. Confirming Harvey’s finding [6], the processes surrounding the use of such systems seemed to be more important as factors in promoting change than the choice of system.

The framework used for the analysis was developed from the findings of this study and from the evaluation literature which provides detailed accounts of evaluation processes in operation and examples of procedures which have been developed to contribute to the validity of findings. In his classic discussion of metaphors for evaluation, Smith distinguishes between four levels at which evaluation strategies can be usefully compared: the paradigm level, the level of discipline, the operational level and the level of technique [18]. These latter two levels, the operational and the technical, provide us with a useful means of distinguishing between the techniques which are used for audit, and the social processes through which they are mediated. On the whole, audit literature in the UK has concentrated more on the technical issues, for example in its consideration of data collection methods, sampling, and project designs [19,20]. However, it is with regard to operational matters that it becomes possible to discuss a range of different issues which concern the possible roles, tasks and procedures for staff and other stakeholders to take in audit projects. This allows analysis not just of the types of data collected, for example, but of who collects them and how this is done. These considerations of processes have been largely neglected, yet they need to be analysed not only for their contribution in terms of project design, but also for their ability to involve practitioners appropriately in audit projects. Process considerations are highly significant to the success of audit projects, yet they have generally been treated as an issue.
were obliged to solve. Therefore the quality and scope of procedures were insufficiently well specified, thus posing variables. Thirdly, all systems need instructional support for decision-making problems which individual auditors and their position as 'insiders' or 'outsiders' [23], and the nature of staff involvement in the audit process are crucial. Neglecting such issues runs the risk that findings may be ignored unless the question of how organizational relationships may obstruct or contribute to the change process is addressed [7]. In this context, the identity of the auditor(s) and their position as 'insiders' or 'outsiders' [23], and the nature of staff involvement in the audit process are crucial variables. Thirdly, all systems need instructional support for their successful operation, and training is increasingly regarded as an essential element in successful audit [24]. One of the key findings of the initial CNPNA evaluation was that procedures were insufficiently well specified, thus posing decision-making problems which individual auditors de facto were obliged to solve. Therefore the quality and scope of advice and guidance available for system users need to be assessed.

The core element of an audit system is the scheme of topics or standards which are audited. We deemed this to be beyond the scope of an analysis of audit processes. However, there is an issue about the sources of evidence used to test performance, and the types of tools which are used. We identified four key issues relating to data collection: firstly the application of triangulation principles by using a range of different sources. Secondly, we followed the evaluation literature [25,26] in recognizing the status of audit data as potentially contestable and their meanings open to negotiation, and therefore included a category to cover the nature of negotiability permitted within the system. Finally, the format (e.g. the use of scales, open-ended questions, etc.) which is used for recording data, and the recommended sampling methods, are important elements in the overall approach. A fifth and final category covers the way in which audit results are handled to facilitate completion of the audit cycle, and incorporates factors which are known to influence the receptivity of staff towards engaging with the results of an audit exercise and developing local action plans for quality improvement. Thus, the provision of a framework for dissemination of findings and the implementation of resulting improvements is a basic system requirement [27,28] and the status of audit findings in terms of their ownership is a key issue given detailed consideration in the evaluation literature [7]. The way in which results are analysed and presented allows us to evaluate the overall scheme, and the ease with which raw data can be consulted indicates the potential for specific problems identified in the audit results to be further illuminated.

These five categories can thus be summarized as follows:

- the status of the system, including any claims (to applicability and to validity);
- the constitution of the system within the organization;
- the provision of advice and guidance for operating the system;
- the approach adopted within the system towards data collection;
- how the system recommends that results are handled.

The six systems compared

The evidence presented below is based on all available documents associated with each of the systems, and has been confirmed as factually accurate by their authors.

Status of the system

The status of the system refers to its applicability and accessibility from the end-user's viewpoint, including the existence of facilities for local customization and review of the system itself, any evidence quoted for its validity, and the existence of any procedures for establishing inter-rater reliability (Table 1).

In relation to local customization, we can distinguish two main approaches: either the system itself may be a local one offered for others to use or adapt in unspecified ways, or it may aspire to universal applicability. In the latter case, authors either recommend that minor editorial changes may be made to the system's contents, or that local auditors may select particular sections of the system to work on. Of the systems (Achievable Standards of Care and QUEST) are explicitly local, although both express the hope that they might be considered useful in other areas. The CNPNA and the paper-based version of PNM allow for minor changes to their content, while SCAPA, QUARTZ and the computer-based version of PNM allow the strategy of selecting a specific section to begin with. These latter two allow questions to be added, and QUARTZ permits local values and aims to be defined.

Only QUARTZ and Achievable Standards of Care refer to the possibility of reviewing the system itself after use. In the former case, an evaluation pack is included in the system, and in the latter, the intention is expressed for the system's authors to redefine standards when those set out in the system have been achieved throughout the Health Authority. By default, three of the four other systems imply that the standards they define are universally applicable (with minor local modifications), while QUEST distinguishes some which specifically apply to its local Health Board.

For those systems which claim their contents are universally applicable, the issues of validity and inter-rater reliability are crucial, since unless these can be established there is no support for the claim. Although the question of what might
Table I  Status of the system

<table>
<thead>
<tr>
<th></th>
<th>CNPNA</th>
<th>SCAPA</th>
<th>QUARTZ</th>
<th>ASC</th>
<th>QUEST</th>
<th>PMN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Price¹</td>
<td>B</td>
<td>B</td>
<td>C</td>
<td>B</td>
<td>A</td>
<td>B (paper system)</td>
</tr>
<tr>
<td>Local customization</td>
<td>Minor</td>
<td>Omissions permitted</td>
<td>Extensive</td>
<td>None</td>
<td>None</td>
<td>Paper system: minor; Computer: extensive</td>
</tr>
<tr>
<td>Evidence of validation</td>
<td>None</td>
<td>Reference to pilots</td>
<td>Published evaluation</td>
<td>None</td>
<td>Reference to field testing</td>
<td>Reference to trials</td>
</tr>
<tr>
<td>Procedures for inter-rater reliability</td>
<td>Provided</td>
<td>NA</td>
<td>NA</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Facility for review of instrument</td>
<td>None</td>
<td>None</td>
<td>Provided</td>
<td>Intention expressed</td>
<td>None</td>
<td>Paper system: none; Computer: possible</td>
</tr>
</tbody>
</table>

¹ A, unpriced; B, under £50; C, £50–100. NA, not applicable.

Count as a test of validity for an audit or QA system is both complex and disputed [5], in fact only one of the systems quotes documented evidence to support a validity claim of any kind. Two systems, the CNPNA and Achievable Standards of Care (which is an explicitly local system anyway) make no reference to the question of validity. Of the other four, SCAPA, QUEST, and PNM make reference to unpublished and therefore unverifiable pilot testing, while QUARTZ refers to an extensive, published in situ evaluation [29]. PNM also makes what must be regarded as a secondary claim, that it is based on the more thoroughly tested Rush-Medicus system [30].

Regarding inter-rater reliability, only the CNPNA provides auditors with advice on establishing this through the use of trial runs. Five of the systems, however, provide advice on how operators should interpret the rating scales provided.

**Constitution of the system**

The constitution of the system covers questions about who operates the system, and how their relationships are constituted within the wider organization, its management structures, its professional forums, and not least, those personnel working in the setting under scrutiny. This feature relates to issues which are critical aspects of evaluative activities but which are often overlooked because an expert operator of the system is assumed to exist at the interface between the system and the organization.

The six systems have been analysed in relation to three levels: the way the system links with the wider organization; the identity of the auditor(s); and the nature of involvement of staff in the setting being audited (Table 2).

Four of the six systems do not discuss the way in which the system relates to the organization in any respect. Of the other two, QUARTZ gives detailed guidance about review processes, and PNM recommends the convening of a special project group for periodic administration of the system and selection of auditors. Thus advice concerning decisions about who will operate the system is absent in four out of the six systems.

Five of the six systems, however, specify the positions of personnel required to operate them, and three models are apparent. Two systems recommend a single person at ward manager level (as a minimum), internal to the hospital and external to the ward, and although it makes no specification, SCAPA implies that this model is recommended. QUARTZ recommends this model too, but allows an alternative strategy of convening a multidisciplinary audit team. This is the model proposed in Achievable Standards of Care where the composition of the multi-disciplinary review team is specified. PNM lies between these two models, using a pair of auditors both external to the ward.

The nature of involvement on the part of ward staff also varies. Two of the systems recommend preparation sessions with staff. All of the systems make use of staff as the subjects of data collection, but only two actively involve staff in discussion as part of this process. Only Achievable Standards of Care includes staff members (the charge nurse and the physician) in an audit team. PNM allows staff to suggest amendments to the contents of the audit schedule.

**Provision of advice and guidance**

The extent to which an audit system provides advice and guidance for operators can be crucial to its smooth implementation. Any paper- or computer-based system requires basic informational support about strategies for its operation, avoidance of errors, and the means of solving problems which arise. In the case of these six systems, four different types of advice and guidance are distinguished for the purposes of analysis: training requirements to operate the system; estimates of the amount of input in terms of the time required.
Audit systems for mental health nursing

Table 2 Constitution of the system

<table>
<thead>
<tr>
<th></th>
<th>CNPNA</th>
<th>SCAPA</th>
<th>QUARTZ</th>
<th>ASC</th>
<th>QUEST</th>
<th>PNM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wider organization</td>
<td>No reference</td>
<td>No reference</td>
<td>Located within organization</td>
<td>No reference</td>
<td>No reference</td>
<td>Project group</td>
</tr>
<tr>
<td>Auditor(s) identity</td>
<td>Individual: outside setting</td>
<td>Not specified</td>
<td>Individual or team outside setting</td>
<td>Team: inside and outside setting</td>
<td>Individual: outside setting</td>
<td>Pair: outside setting</td>
</tr>
<tr>
<td>Nature of staff involvement</td>
<td>Data source data collectors</td>
<td>Data source discussion</td>
<td>Data source discussion</td>
<td>Data source audit team member</td>
<td>Data source preparation</td>
<td>Data source preparation</td>
</tr>
</tbody>
</table>

Table 3 Provision of advice and guidance

<table>
<thead>
<tr>
<th></th>
<th>CNPNA</th>
<th>SCAPA</th>
<th>QUARTZ</th>
<th>ASC</th>
<th>QUEST</th>
<th>PNM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training required</td>
<td>Through trial runs; available from authors</td>
<td>None specified</td>
<td>Workshops available</td>
<td>None specified</td>
<td>Importance stressed</td>
<td>Through trial runs</td>
</tr>
<tr>
<td>Estimate of inputs</td>
<td>Provided for data collection</td>
<td>Local agreement</td>
<td>Provided in detail</td>
<td>Not provided</td>
<td>Some provided</td>
<td>Some provided</td>
</tr>
<tr>
<td>General guidance</td>
<td>Minimal</td>
<td>None</td>
<td>Detailed</td>
<td>None</td>
<td>Minimal</td>
<td>Minimal</td>
</tr>
<tr>
<td>Guidance on data collection</td>
<td>Quality scale: minimal</td>
<td>Specific for each question</td>
<td>Detailed: manual provided</td>
<td>Specific for each question</td>
<td>Scoring</td>
<td>Scoring</td>
</tr>
</tbody>
</table>

Two of the systems provide additional training for operators beyond the materials in accompanying manuals, one of which offers a range of consultancy services including a formal 3-day training course. The level of skill needed to operate the system is explicitly identified by only two of the systems, while one notes the importance of educating staff in general about the system. One further training requirement is raised in two of the systems, and that is the need for operators to carry out ‘trial runs’ of aspects of the system. In two cases there is no mention of any training needed.

All systems but one address the question of how much time will be involved in conducting an audit. They vary, however, over the level of detail provided, ranging from SCAPA which simply specifies the need for an agreement to be made between the auditor and the staff, to QUARTZ which provides detailed timetables and accompanying estimates for reviewer and staff. The CNPNA is highly specific in relation to its activity analysis section, for which data collection involves time-sampling techniques.

In relation to the provision of general guidance on auditing, the contrast between QUARTZ and the five other systems is very marked. In two there is no guidance; in the remaining three there are a few sentences with the addition of pro forma letters of invitation in one, but for QUARTZ there are two manuals of approximately 100 A5 pages each. This and other similar process-related discrepancies may arise from the fact that QUARTZ identifies itself as a QA system, and hence more explicitly a system of management than the other five systems which describe themselves as audits or review.

The one area in which all of these same five audits do provide guidance is in the use of their rating scales. All pay some attention to explaining the meaning of the scales and how to score them, although as indicated earlier, only the CNPNA addresses the associated question of establishing inter-rater reliability. Achievable Standards of Care devotes an entire section to guidance on how to rate four criteria, and SCAPA gives individual guidelines in relation to many of the statements which are to be rated. In the case of QUARTZ, a different approach is adopted in which the reviewer is required to make ratings in relation to the stated values of the care setting. Rating scales, it should be noted, are a technique which allows measurement to take place, and the level of attention paid to the ‘correct’ use of these scales indicates the high value placed on measurement within these five systems.
Table 4 Approaches towards data collection

<table>
<thead>
<tr>
<th></th>
<th>CNPNA</th>
<th>SCAPA</th>
<th>QUARTZ</th>
<th>ASC</th>
<th>QUEST</th>
<th>PNM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range of</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Not specified</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>sources</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negotiability</td>
<td>No reference</td>
<td>Discussion</td>
<td>Discussion</td>
<td>Discussion within team</td>
<td>Prior to data collection</td>
<td>No reference</td>
</tr>
<tr>
<td>Format</td>
<td>Scores</td>
<td>Scores</td>
<td>Comments</td>
<td>Open-ended questions</td>
<td>Scores</td>
<td>Scores</td>
</tr>
<tr>
<td>Samples</td>
<td>Specified</td>
<td>Not specified</td>
<td>Reviewer’s decision</td>
<td>Not specified</td>
<td>Not specified</td>
<td>Specified</td>
</tr>
</tbody>
</table>

Approach towards data collection

An episode of data collection is an essential part of any audit or QA system. However, the types of data which are deemed useful for an audit will vary not only according to the subject of the audit, but also to the overall approach taken. The six systems were analysed according to four different characteristics related to data collection: the range of data sources used, the extent to which the system permitted the meaning of data to be negotiable, the formats used for data collection, and any recommendations regarding sampling (Table 4).

All of the systems except one require the use of multiple data sources. Two make specific reference to the necessity for gathering corroborative evidence from additional sources. A similar range of data sources is recommended in all cases - staff, patients, documentary evidence, and observation by auditors. One system, QUARTZ, additionally recommends the use of group discussions for data collection. It should be noted that the system which does not explicitly recommend this technique (Achievable Standards of Care) nevertheless uses a different but similarly powerful mechanism to attempt to prevent distortion and bias – the collection of data by multiple operators in the form of a multi-disciplinary review team.

The issue of negotiability of data is complex and important. For a practice-oriented evaluation-like audit, some means must be provided whereby those in the setting under audit may, if necessary, register a different view about the meaning of the data gathered. Without such provision, practitioners may choose to reject findings on a whole range of different grounds, from the basic accuracy of the data to its relevance and general applicability. Two of the systems make no reference to such a facility. The other four systems vary regarding the points in the audit process at which negotiation might be possible. In QUEST, there is an initial interview with the charge nurse which provides an opportunity for amending the questions through which data are collected, while in SCAPA findings are discussed between ward staff and auditor before any action is agreed. In Achievable Standards of Care, discussion of findings is limited to members of the review team (who, it will be recalled, are drawn from both inside and outside the setting), while for QUARTZ negotiability exists both during data collection for some schedules – especially those which use group discussion – and during feedback for all schedules.

Regarding the form in which data are collected, all systems except two use more than one approach to the coding of data. Every system uses one or both of the following two formats: yes/no/not applicable (with some allowance for half marks), and four- or five-point rating scales. The CNPNA includes a grid for time-sampling of nursing activities. Only two systems make provision for the inclusion of qualitative information, QUARTZ in the form of open-ended questions within some of the schedules, and SCAPA by allowing for comments to be recorded for every scored statement.

Only two of the systems make recommendations about the scale of data collection – e.g. how many subjects should be interviewed – and sampling techniques, and both are quite specific about the size and type of sample to be used. In none of the other systems is any reference made to this except for QUARTZ which requires the reviewer to make decisions about sampling.

How results are handled

The way in which results or findings are handled reflects to some extent the type of data gathered. However, there are also some important processes at stake in this phase of the evaluation process. With regard to the analysis procedures (Table 5) we distinguished six different characteristics: how the data are analysed, structures for implementing results, how the results are presented, who owns them, how they are disseminated and whether any facility exists for questioning scores. This latter relates to the previously discussed issue of negotiability of data.

Data analysis and presentation

For two of the systems, PNM and the CNPNA, computerized data-processing facilities are available for the input and analysis of data. In the remaining four, the operator needs to perform calculations manually. Achievable Standards of Care
contains a grid summarizing all the questions into which raw scores can be entered, providing a visually accessible overview of the data. A quite different approach is adopted by QUARTZ, requiring reviewers to write a report which would contain summarized data.

The way in which results are finally expressed is crucial as a vehicle for communicating them not only to the setting under audit but also to other interested parties. Four of the systems opt for percentages, with PNM providing a single global percentage index, QUEST a percentage for each of the scored sections, and the CNPNA providing a global score along with breakdowns into subcategories. SCAPA provides proformas for aggregating results first as percentages for each subsection, and then for dividing these percentages into three global ‘action’ categories. Given the small number of items within the subsections (range 1–14), this process of aggregation and disaggregation runs the risk of distorting data which could just as easily be presented raw.

Visual techniques are employed in three of the systems: Achievable Standards of Care provides a method for presenting the raw data graphically by means of a grid on which raw scores are plotted; the CNPNA presents the results of the nursing activity analysis as a series of bar charts, while SCAPA goes so far as to use imagery by depicting the final three ‘action’ categories as sun, clouds and rain.

Again, QUARTZ is an exception to the other five systems in that there is no attempt to quantify or measure through the use of scored dimensions. Instead the reviewer writes a report which may nevertheless contain quantified information, but which is structured according to the strengths and weaknesses of the setting.

Ownership, implementation and dissemination of results

The ownership of audit findings is a problematic issue. Differing protocols have been developed at local levels, and in some cases of medical audit the principle of confidentiality — that ownership should rest solely with the contributing practitioners — has proved an obstacle to the publication of results because of the scale of the consultation processes required for clearance from a large number of contributors.

The two issues of ownership and dissemination are very closely related, as agreements about one will influence possibilities for the other. Only three of the systems explicitly address the question of the locus of ownership of the audit findings, and all agree that these should be the property of the setting under audit. Neither the CNPNA nor Achievable Standards of Care refers to this, although the latter does suggest a dissemination procedure. The sixth system does not explicitly address the issue but seems to imply that findings belong to the senior nurse and the charge nurse who would have conducted the audit.

Wider dissemination of audit findings are discussed in all systems but one. All take the form of a report. In the case of SCAPA, distribution appears to lie in the hands of the ward manager, whereas in QUARTZ the final quality report is written by staff and is given in the first instance to immediate line management and any further dissemination is agreed through consultation between reviewer and staff. Two systems advise that one copy should be left on the ward and one copy made for senior managers, while one simply advises that a copy be given to senior managers.

It must be noted that unless the issues of ownership and dissemination are addressed within a system, then one thing is certain — decisions about these matters de facto become the business of those who are most familiar with the audit process (often the operators), who may have neither the power nor the status within the organization to take such decisions. The likely outcome of this scenario is either that decisions are avoided, or that they are taken without appropriately full consultation.

Similar arguments apply to the question of how audit
findings may be implemented. In the absence of clear understandings about ownership and dissemination, the business of implementing any recommendations arising from the audit again rests de facto with staff in the setting under audit who may or may not agree with the findings (see the section in which negotiability of data is discussed). One system makes no reference to how findings could be implemented while three suggest the drawing up of action plans with review dates. QUARTZ involves the staff in the setting under audit in this process by adding an extra stage where they produce their own quality report as a precursor to action planning. In two cases the ward manager and senior manager or quality co-ordinator are advised to engage in further discussions on implementing changes resulting from audit findings.

One very helpful facility which supports the implementation of results is the ability to interrogate aggregated data so that where problems appear to exist, action can be informed by reference to the relevant raw data. Achievable Standards of Care allows the most accessible route to the raw data via the grids mentioned previously. In all other cases, ease of access depends on how the raw data are presented in the audit documents. With the CNPNA, the volume of calculation required is large enough to prohibit scrutiny of raw data on anything other than an exceptional basis. PNM, however, has produced a computer software package which actively encourages operators to intelligently disaggregate the global scores and discover which questions or groups of questions contribute to low scores.

Discussion

As the foregoing analysis shows, the extent to which each of these six systems addresses issues related to processes varies considerably. The system which is least forthcoming is the CNPNA, and the one which goes into greatest detail is QUARTZ. Of the others, two vary in ways which point on the one hand to an affinity between the paper-based version of PNM and the CNPNA, and on the other between SCAPA and QUARTZ. The remaining two systems, it should be noted, both originated as local systems which are essentially being offered to a wider audience to facilitate the development of other local applications.

These differences might usefully be conceptualized as the way in which tools can be distinguished from instruments, where tools have close relationships with their objects (perhaps to repair or reconstruct them) and instruments are more distant, their function being to inform. This takes further an analogy proposed by Carter et al. in relation to measures of performance in distinguishing ‘dials’ from ‘can-openers’ where the authors’ aim was to compare different approaches to the presentation and use of information – whereas dials only display scores, can-openers allow further examination of data [31]. Like all analogies, such ideas should perhaps not be taken too far, but they do help in thinking about how well a system specifies the way it might engage with the object under scrutiny.

The above analysis shows the CNPNA and the paper-based version of PNM as instrument-like systems, which aspire to the status of measuring devices applying equally well to different settings because they purport to generate scores along a single dimension, or at least an aggregate of dimensions. Hence they produce percentage scores on the basis of data in fixed categories collected by dispassionate detached operators using techniques of ‘random’ sampling. This very detachment militates against close engagement with the setting under audit, which is why the two emerge as relatively unforthcoming.

In contrast, QUARTZ and SCAPA are more like tools because they not only acknowledge but utilize the relationship between the setting under audit and the operators of the system, for example by using group discussions between staff and operator, through allowing negotiation in the interpretation of data, and by the incorporation of action planning in the system. In so doing they reject the mantle of objectivity to which the two previously described instruments aspire, while capitalizing on their potential for stimulating action within the setting by involving staff in the processes of evaluation. This rejection of objectivity is more complete in the case of QUARTZ where no attempt is made to produce the canonical feature of an objective system, the scaled measurement. SCAPA, in contrast, does make this attempt but is unsuccessful because the percentages it produces are no more useful than the raw data.

Conclusion

Our purpose in developing this framework for comparing different audit systems was for the development of a new one which would improve not only on the contents (i.e. the topics) of existing systems, but also to improve on existing audit practice. We needed to consider which of these two general approaches represented a suitable paradigm for developing any new system with wide applicability (the two local systems being ruled out as a paradigm because they are local). In addition, we needed to consider the changing policy context which has encouraged a move from uni-disciplinary to multi-disciplinary audit.

The case for producing a measuring device is quite strong in so far as validity and objectivity are features highly valued by many managers and professionals in the UK [32]. However, there are several crucial arguments against it. Firstly, as we have seen above, the systems we have analysed actually fail, in their own terms, to deliver the kind of validity they require within the paradigm. They are constructed around an assumption, so far not borne out in general in the literature [33,34], that quality of care can be reduced to one, or more than one, dimension. Nor do they make any claims to validity by reference to any empirical evidence in their own case regarding the existence of such dimension(s). Furthermore, they quote no evidence to support the use of measuring scales divided according to percentage scores which ought to have equal meaning along the entire scale. For the CNPNA, one of the authors is currently engaged in attempting to
establish normative values for aspects of the system from data gathered by operators in several hospitals (K. Hurst, Nuffield Institute for Health, Leeds, personal communication). While this represents an attempt to develop validity through norm-referenced criteria, it nevertheless has the consequence of reducing current users of the system to the status of pilot studies.

Secondly, there is the argument for which the evidence is rehearsed above – that instrument-like systems by definition must distance themselves from the setting under audit until the point at which a final report is produced, thereby severely reducing the opportunities for staff, clients and their carers to become involved in the process. Given that audit and QA are explicitly action-oriented systems, this is an important disadvantage which might only be worth tolerating if the final results could claim the kind of indisputable validity a truly objective measure would produce.

Thirdly, the move from uni-disciplinary to multi-disciplinary audit required us to consider the different approaches to audit among the various professions. In the UK, medical audit has developed in the direction of exploring single topics in depth [38], in marked contrast to the traditional nursing approach of auditing according to pre-determined topics offering a comprehensive description of the care process. Since successful collaboration among disciplines is likely to decrease the more divergent are their approaches to audit, the instrument-like systems seemed less suitable candidates for developing such collaboration. Furthermore, since tool-like systems set out more clearly the nature of the audit process, they also offer concrete opportunities for collaborative audit project management and implementation, such as the ad hoc multi-disciplinary audit teams of Achievable Standards of Care.

Finally, there is a debate about the feasibility of producing national standards in relation to aspects of the quality of care. This is a difficult and contested area in the UK NHS [36,37] and there are problems about defining and agreeing what kind of features of care are amenable to treatment of this sort. There has been an ongoing debate throughout the public sector in the UK about national systems of performance monitoring since the introduction of the idea through the Treasury’s Financial Management Initiative in 1981 [38]. Currently it is most hotly debated in relation to such diverse public services as the performance of hospitals, schools and local police forces, for which government wishes to produce ‘league tables’ of comparative indicators on activity and outputs.

However, care settings vary not only in relation to the resources they have to offer, but local circumstances may also cause practice itself to change by developing more finely-tuned criteria to underpin clinical standards, or through new developments in clinical practice. Nevertheless, the wider use of the two locally-devised systems, Achievable Standards of Care and QUEST, testifies to the potential for the wider applicability of sets of standards devised in one setting which could be transferred to others. The Patient’s Charter – part of a wider Citizen’s Charter initiative introduced in the early 1990s [39] – has also shown that it is possible to devise standards with a strong clinical component which can apply nationally, some of which are beginning to be used in national ‘league tables’. But the Charter also recognizes that there is a need for some flexibility at local level. The debate about national standards, however, has yet further implications for instrument-like systems, for they too set out to define standards as universally applicable. In the present policy context where quality initiatives such as the Patient’s Charter and local standard-setting activities are developing fast, there is considerable potential for conflict amongst local audit operators and practitioners about which standards to follow. The case therefore is very strong for devising a system which seeks to integrate these different quality initiatives rather than to introduce a further dimension claiming to be authoritative.

The framework utilized in this paper has allowed us to address the problem, largely neglected by others, of suitable processes for carrying out audits, and how these processes can be analysed so that they may contribute towards the validity and rigour of the system. The framework has also shown that by widening the scope for analysing audit, so that practices and contexts are acknowledged as being socially constructed rather than socially given, these social dimensions can be treated as processes which are amenable to explicit instead of implicit manipulation. Including these questions within the frame of discourse also widens the discussion about strategies for ensuring methodological rigour, so that process issues can be utilized to enhance the validity of audit findings, rather than treated as influences at best to be controlled and at worst to be excluded altogether.

A framework with this breadth of scope also offers a means for analysing and comparing a wide variety of audit projects, and should therefore prove useful to all stakeholders in the audit process – including the recipients of care [40] – who need to perform judgements about the merit of different projects.

Acknowledgements

This study was supported by funding from the NHS executive and the Northern Regional Health Authority.

References


Received in revised form 3 September 1997