Critique of the medical record is a time-honored approach in the evaluation of trainees. Most faculty members have some experience evaluating student write-ups as part of their clinical clerkships. However, systematic review of the medical record is done much less frequently in residency and fellowship education. Medical records serve a number of important functions: (1) an archive of important patient medical information for use by other health care providers and patients; (2) a source of data to assess performance in practice such as treatment of specific chronic medical conditions (e.g., diabetes), postoperative care, or preventive services; and (3) the documentation of clinical decisions. One can readily see how these patient care functions of the medical record can be used for educational and evaluative purposes.\(^1\)

In the United States, the Residency Review Committee (RRC) of the Accreditation Council for Graduate Medical Education (ACGME), the organization responsible for the accreditation of training programs, requires medical record audits as part of the training program’s evaluation.\(^2\) The accreditation requirements specifically state that the program director should ensure that a representative sample of trainees’ medical records are audited for quality of documentation and information during resident inpatient and outpatient clinical rotations, with feedback given to residents. Medical record audits are an essential element in the evaluation of the new competency of practice-based learning and improvement (PBLI). PBLI requires that residents be actively involved in monitoring their own clinical practice and improving the
quality of care based on a systematic review of the care they provide. The American Association of Medical Colleges also endorses the importance of skills in medical records for medical students. Chapter 11 covers how and why the data from medical record audits are essential to evaluate a trainee’s competence and performance in quality improvement. Medical records can also be used to assess clinical reasoning through a technique known as chart-stimulated recall, described in the section Chart-Stimulated Recall. Both activities promote self-reflection, an important skill needed for lifelong learning.

Sources of Data for Practice Audits

In this new era of information technology (IT), data are often available from other sources besides the “paper-based” medical record. Many hospitals, and to a much lesser extent outpatient clinics, are moving to electronic medical records to document visit encounters. However, the “written” medical record remains a vital component of the educational experience, whether in paper or electronic form. Other potential data sources for audit include computerized laboratory data and radiographic records, claims and pharmacy data, and other administrative databases. For example, quality improvement organizations (QIOs) use Medicare claims data to track the quality of care for Medicare beneficiaries who receive outpatient care for chronic conditions such as diabetes (eye examination rates, hemoglobin A1c testing), prevention screening tests (mammography, colonoscopy), and immunizations. Programs can work with their state QIO or health plans to obtain aggregate data on outpatient performance. Each particular type of data system has its own set of limitations, so it is important to ask your local QIO and IT department what types of data are available at your institution in your specific specialty.

Paper-Based Medical Records

The paper-based written medical record is still the most common format used to document clinical care activities and can provide valuable data to evaluate and provide feedback about the “quality” of care. However, most experts believe that there should be more rapid uptake of electronic medical records (EMRs); current data suggest only 10% to 15% of outpatient practices in the United States have fully operational EMRs. Until EMRs are more widely adopted, medical educators will need to understand how to use and extract important information from paper-based medical records for education and evaluation. Audet showed that less than 30% of U.S. physicians are using any performance data to improve their care practices; our trainees must be better prepared.

The major limitation of the paper-based, and even electronic, record is that the record is only as good as the information contained in it. First, research has shown that important aspects of the clinical encounter are often not documented (see Potential Disadvantages of Medical Record Audits), and the quality of the written information is highly variable from trainee to trainee (assuming, of course, you can read it!). Second, paper sheets are like socks in a dryer; they tend to get lost all too easily. Third, paper records require an inordinate amount of personnel time to maintain. These aspects should be considered when deciding how to use paper-based records in your evaluation system.

The Electronic Medical Record

For years the term “medical record” referred to the collection of written information, including history and physical examinations, laboratory and radiology results, problem lists, and so on, contained in the patient’s paper chart or file.

The introduction of EMRs is beginning to substantially alter the way patient clinical information is organized and used for the delivery of medical care. As a result, the EMR can be expected to also alter the way in which we use the medical record for evaluation. The effect of computer-based record systems on documentation in training programs is not well known, but deserves further study as many institutions move to electronic records. To date, the main effect of EMRs has occurred in the inpatient setting, with many, but not all, studies showing improvement in patient safety and a reduction in medical errors. A number of these EMRs provide the mechanism to enter the medical history and physical examination. Little is known on how EMRs affect the quality and nature of the medical trainee’s documentation practices. However, electronic records may be highly valuable in helping to determine the actual clinical experiences of trainees. One study at a single hospital in Boston with a computer-based records system found that more information per each patient’s problem was being entered into the computer record, as opposed to a paper record. Although this study did not address the direct impact on assessment of competence through electronic record review, the study does suggest that computer records may provide a greater quantity of information about a patient encounter over the written medical record.

However, we have noted one serious documentation problem we call the “cut and paste” syndrome, in which trainees cut and paste previous notes for use in admission and daily progress notes, with or without adequate editing. In one of my previous hospital’s internal quality improvement activities, we noted this was a common activity for “efficiently” completing daily progress notes. However, unless we’ve moved into a new time dimension, I found it hard to believe a patient was “postoperation day 1” for 7 consecutive days. In another review of inpatient electronic charts for a quality of care project on pneumonia, we uncovered a number of instances in which the “cut and pasted” information by trainees was both erroneous and failed to account for changes in a patient’s clinical status. Educators should be particularly sensitive to this practice.
EMRs can potentially make the retrieval of specific types of clinical data for review much easier. Unfortunately, pulling data at the individual practitioner level for specific categories of patients is very difficult at this time for most inpatient and outpatient EMR products. This “registry” function is important for several reasons. First, clinical data are necessary for performance assessments and quality improvement projects. Without robust clinical data targeted for specific populations of patients (e.g., those needing preventive services, patients with chronic disease such as diabetes), it is almost impossible to implement quality improvement. Second, trainees need to know what types of patients and conditions they are seeing in clinical practice. At the current time few programs have the capacity to track the clinical experiences of trainees. For educators who have EMRs as part of the clinical environment, we recommend approaching the information technology or quality improvement departments of your organizations to see what types of information can be retrieved from the EMR system for trainees.

**Claims Data**

At the residency and fellowship levels, trainees and office staff routinely use ICD and CPT codes for patient visits, especially in the outpatient setting. This information can be a valuable source of information about the clinical practice of trainees. For example, claims data can be very helpful in identifying a cohort of patients. We used the claims database at National Naval Medical Center to identify the population of all diabetic patients seen in an internal medicine residency clinic. Using this data, we were able to then access the laboratory database to assess processes of care (measurement of hemoglobin A1c, lipids, microalbumin) and the degree of glycemic control (hemoglobin A1c level) and hyperlipidemia for each trainee’s panel of diabetic patients. We also used this claims database to “track” the make-up of each resident’s patient panel. Likewise, the claims database can be used to identify a group of patients admitted with conditions such as acute myocardial infarction (AMI) and pneumonia. This can then facilitate the pulling of charts for review. Because AMI and pneumonia are target conditions for Medicare, your quality improvement department may already be reviewing the quality of care for these groups of patients.

Several caveats should be noted about using claims data for evaluation. First, the use of claims data to measure “quality” is highly dependent on the quality of the coding. Poor coding practices can limit the value of the claims data. Second, claims data are essentially limited to the process of care and usually cannot provide specific detail about the care received.

**Laboratory and Other Clinical Databases**

For most hospitals, but to a much lesser extent for outpatient settings, the laboratory, pharmacy, and radiology data for groups of patients are available electronically. Access to this type of electronic data greatly facilitates the systematic review of chronic illness care, cancer screening, etc. Unfortunately, for many training programs, patients often receive services from multiple locations, making it more difficult to track these services. Mammography is one example of a test for which the patient may have several options concerning where it is performed (offsite office, mobile van, hospital, etc.). In this situation, those using only the local hospital database are likely to significantly underestimate the use and receipt of certain services.

**The Audit Process**

Understanding the basics of the audit process is crucial to maximizing the utility of medical records as a tool for both formative and summative assessment. Because medical record audits can be time-consuming, you should not perform an audit until you are clear about the educational and evaluation purpose of the audit. The audit cycle is closely related to the PDSA (plan-do-study-act) quality improvement cycle developed by Shewhart over 60 years ago. The audit cycle (Fig. 5-1) highlights how information from a medical record audit can help trainees to improve and progress professionally.

The simple diagram in Figure 5-1 highlights the importance of clinical practice data as a catalyst for individual change. Without such data, it is nearly impossible to determine “quality” of performance and to measure progress. As we will see in Chapter 11 on practice-based learning and improvement and systems-based practice, data from medical record audits form an essential component. Trainees must not only understand this simple audit cycle, they must have the opportunity during training to perform all of its steps.

The value of the audit process is only as good as the information abstracted from the medical record. There are two main approaches to the conduct of an audit: “explicit” and “implicit” review. For years, the most common approach was implicit review. Implicit means the auditor does not have strictly defined criteria when reviewing a medical record. Instead, the reviewer relies on general guidelines to determine if care delivered, based on the medical record, was “good” or “bad.” Implicit review was commonly used in the...
Advantages of the Medical Record Audit

As an evaluation tool, audit of the medical record has a number of important strengths. Some form of medical record audit should be part of every training program’s evaluation system. The specific advantages of medical record audit are described here.

Availability

Medical records or other clinical data are usually available and accessible. Getting to the record is usually not a major problem, but depending on the type of record (paper, electronic), pulling out specific aspects of care can be a challenge. Electronic patient registries are best for creating population-based reports for specific quality process measures, but even the use of flowsheets and problem lists can greatly facilitate the collection and analysis of the quality of chronic care, acute care, and preventive services.

Feedback

Medical records allow for corrective feedback centered on actual clinical care in a timely manner. Most faculty are required to review, and often co-sign, trainee notes on clinical rotations. Too often, faculty fail to take advantage of information obtained from the medical record for use in their evaluation of and feedback to the trainee. In fact, the written medical record can be used as a “guide” to query the resident about the choices of specific diagnostic or therapeutic approaches for the patient. This approach is known as chart-stimulated recall (CSR) and is discussed in greater detail later in the chapter.

Changing Clinical Behavior

The majority of studies have shown that chart review can change trainee behavior through direct feedback such as that provided by “report cards” on performance of targeted clinical interventions (e.g., prevention measures).

Early studies demonstrated audits could be very effective when explicit criteria were used for data abstraction. Martin and associates reported in 1980 that a group of residents subjected to continuous chart review demonstrated a 47% reduction in laboratory usage in comparison to a control group of residents receiving no review. In fact, chart review with feedback was more effective than a “financial incentive” (a textbook). Kern and associates found that a chart review combined with feedback improved performance in record documentation and compliance with preventive care measures. Several other studies found that a structured chart review using explicit criteria coupled with written and verbal feedback led to substantial improvement in the delivery of three preventive health measures. One study found that the audit of just three preventive care interventions was associated with substantial improvement in 1980s for patients experiencing critical incidents or adverse events, to review complaints, or for routine peer review activities. In medical education, implicit review is a common technique for judging trainees’ patient “write-ups” on clerkships or other clinical experiences. There are several important limitations to implicit review. First, there tends to be unacceptable inter-rater variation, resulting in low levels of reliability. Second, in the absence of reasonable and consistently applied criteria and standards, it is difficult for a reviewer to determine what constitutes good and bad care, especially for complex cases. Attempts to train reviewers to improve the quality of the implicit review approach have been mostly unsuccessful.

Therefore, to determine what constitutes a high-quality trainee write-up; some basic structure with a minimal set of predetermined criteria is essential. In contrast, an explicit review approach uses detailed criteria to perform a medical record audit. Explicit review is now the preferred approach among most quality organizations. In explicit reviews, the quality measures are carefully chosen and defined to be sure they (1) can be measured with reliability and accuracy; (2) are generalized across clinical sites; and (3) can be aggregated for populations of patients. Likewise, the audit process is also carefully described with well-defined inclusion and exclusion criteria. Box 5-1 provides an example of a quality measure from the U.S. National Quality Forum (NQF), an organization working to standardize a set of quality measures for use by all interested organizations.

For high-stakes decisions about quality of care delivered to groups of patients, the explicit review process is the current “gold standard.”

**BOX 5-1 Diabetes Process of Care Quality Measure**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Percentage of patients with one or more hemoglobin A1c tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numerator</td>
<td>One or more hemoglobin A1c tests conducted during the measurement year identified either by the appropriate CPT code, or at a minimum, documentation in the medical record must include a note indicating the date on which the Hgb A1c was performed and the result</td>
</tr>
<tr>
<td>Denominator</td>
<td>A systematic sample of patients age 18–75 years who had a diagnosis of either type 1 or 2 diabetes</td>
</tr>
<tr>
<td>Exclusions</td>
<td>Exclude patients with a history of polycystic ovaries, gestational diabetes, or steroid-induced diabetes during the measurement year</td>
</tr>
<tr>
<td>Data Source</td>
<td>Visit, laboratory, or pharmacy encounter data or claims; electronic data may be supplemented by medical record data</td>
</tr>
</tbody>
</table>

*From the U.S. National Quality Forum.*
the delivery of six other nonaudited preventive care interventions ("spill-over" effect).26

A recent systematic review by Veloski and colleagues investigating the effects of audit and feedback found positive results in the training setting. Specifically, Veloski reviewed 29 studies that involved residents or a mixture of faculty and residents. Of the 29 studies, 18 (62%) reported positive effects of feedback on clinical care. However, the majority of the studies involved residents at multiple levels, making it impossible for them to comment about the effects of supervision.28

Several other systematic reviews have also found modest positive effects of audit and feedback on clinical care for all developmental stages of physicians.29 Medical record audits appear to be most effective when the data feedback is resident (or physician) specific and the data are provided back to the individual resident for review. Data provided at the group level appear less effective; individuals looking at group data often remark, “I wish my colleagues would do a better job because I know I’m doing better than this!”27

Practicality

Medical record audits allow for a random or targeted selection of patients to be surveyed, and record reviews can be done without the patient physically being present. Furthermore, audits can be scheduled into clinical activities convenient for the training program and resident. We created a half-day rotation in quality improvement for our residents and used a portion of this time for residents to audit their medical records in diabetes and preventive care. Medical record audits are also unobtrusive as an evaluation tool and in this way may help to minimize the “Hawthorne effect.”

Evaluation of Clinical Reasoning

Depending on the quality of the documentation, evaluation of skills in analysis, interpretation, and management is possible. In addition, evaluations of particular patients or conditions can be performed over time, and for many chronic conditions good evidence is available to develop key outcome and process metrics. We’ll come back to this when we discuss chart-stimulated recall.

Reliability and Validity

When explicit criteria are used, a high degree of reliability is possible. This approach applies to such areas as appropriate use of the laboratory, preventive health measures, cost effectiveness, care of chronic illnesses such as diabetes, and the quantity of documentation. Explicit criteria are best suited for process of care measures (such as ordering a hemoglobin A1c on a diabetic patient within a certain time frame) and some outcomes that are easily measured and do not require substantial time (e.g., measuring the level of hemoglobin A1c as a surrogate outcome). Because the information contained in the record relates directly to actual patients, the results of medical record audits have excellent face validity and authenticity. Medical records provide documentation of performance, meaning what a trainee actually does.30 Some studies have also found evidence of construct validity in that results of quality of care audits modestly correlate with cognitive expertise as measured by a secure examination or certification status.31–33

Learning and Evaluating by Doing

Medical record audits allow residents to directly participate in the process of peer review. Egerbretsen commented in 197734 on the positive impact of a peer review system at his family practice residency. Ashton35 made the case for involving residents in hospital quality improvement programs over 10 years ago. Having the residents perform their own audit may be even more powerful; one study involving resident self-audit found the majority of trainees were surprised by results demonstrating they often failed to perform key quality indicators.27 We call this the chagrin, or “a-ha,” factor. A study of practicing physicians participating in a study examining a Web-based self-audit tool found identical reactions from the physicians.36 The main power of self-audit is that the trainee cannot “hide” from the results, and cannot complain about the quality of the data or blame an abstractor for errors because they are the ones who entered most of the data and performed the audit. PGY-2 residents in the Yale Primary Care Internal Medicine Residency Program participated in self-audit as part of a quality improvement experience during their ambulatory block rotations. Residents used part of the time to review their own charts for quality of care in immunizations, cancer screening, and diabetic care. This relatively simple intervention led to meaningful changes in resident behavior and modest improvements in patient care.27 Finally, benchmarking the resident’s performance against some standard, whether internal or external, can also be helpful.26,28,29

Teaching effective medical record audit techniques is becoming increasingly important. Most health insurance companies routinely ask for copies of records to perform reviews of certain practice habits. The Center for Medicare and Medicaid Services (CMS) is also beginning to review the care of Medicare patients in both inpatient and outpatient settings with a goal of public reporting in the near future.37 Accurate and reliable medical record audits are an essential component of many pay-for-performance programs. Thus, involving trainees in the medical record audit process is important for their future success as practicing physicians.

Self-Assessment and Reflection

When the trainee is incorporated into the audit process, the result can be a powerful tool to promote self-assessment and reflection. Given what was just stated about a likely future of public reporting and
continual assessment, physicians-in-training must be prepared to effectively self-assess their own performance, reflect accurately on the results, and then use the results for continuous professional development.\textsuperscript{38,39}

Medical record review can be a very useful educational tool, can potentially change behavior, and can provide useful information when explicit criteria for review are utilized. Such audits can be tracked and included as part of a comprehensive clinical competency record and can be easily incorporated into an evaluation portfolio. Finally, the result of a medical record audit across multiple trainees provides valuable information for program assessment. Audits can identify strengths and weaknesses in the actual care delivered to patients that should play a major role in program assessment and curriculum design. Some would argue that clinical training is only as good as the quality of the care given to the patients. Audits can also be used to assess the effectiveness of educational interventions in the clinical training setting.

### Potential Disadvantages of Medical Record Audits

Despite the tradition of using the medical record as a tool for evaluation of competence, Tugwell and Dok lamented more than 20 years ago over the lack of good research using trainees’ medical records for education and assessment. While the situation is modestly better today, much work remains to be done.\textsuperscript{6} The challenges of the ever-changing and evolving organizational format used for medical records were highlighted in the section Sources of Data for Practice Audits. A few other issues should be highlighted. First, whereas the organizational format used for creating a medical admission or progress note receives a lot of attention in medical school, the same scrutiny seems to evaporate at the residency and fellowship levels. I often felt like I was trying to decipher the DaVinci code when reviewing the progress notes of too many trainees. I’ve seen more modifications of problem-oriented or SOAP (subjective-objective-assessment-plan) notes than I care to remember. Add to this situation the many different EMR vendors all using different organizational formats involving various combinations of templates, checklists, and free text. This new world has created a host of new problems for medical educators. As noted earlier, many educators have discovered trainees often use the “cut and paste” function to simply update progress notes or take information from other notes to complete admission workups. Poorly edited notes electronically copied worsen documentation accuracy and quality, may put patients at risk, and even more disturbing, may represent blatant plagiarism. Despite the obvious efficiencies, I believe the “cut and paste” function should be used sparingly and very carefully.

However, given the clear importance of EMRs for more effective, efficient, and safe health care delivery,\textsuperscript{7} educators must prepare trainees to use EMRs more effectively in the future. This should include what functionality a trainee should look for when implementing an EMR system.\textsuperscript{7} The other challenge is using the medical record as a “measure” of clinical competence. The most important question is, “What are we really measuring about competence in a medical record review?”

### Quality of the Documentation

The quality of a medical record audit can only be as good as the quality of the documentation. Tugwell and Dok\textsuperscript{6} noted, “the fact that records are used more as an aide-de-memoir rather than a documentation of the justification for management decisions, which continues to compromise the validity of the medical record.” This situation may actually be worse today. When trying to assess more than whether certain processes of care were or were not delivered, important questions to ask are as follows:

- Does the record accurately reflect what occurred during the visit?
- Was all pertinent information that was collected during the patient encounter recorded?
- Are impressions and plans justified in the record?
- What facilitating tools (e.g., templates, problem lists, flowcharts) are provided with the medical record?

Do physicians record with completeness what they actually did during the encounter? Norman and associates\textsuperscript{40} in 1985 found, using unannounced standardized patients, that physicians often failed to completely record information obtained and procedures performed. Certain areas, such as physical examination and clinical investigations, were recorded appropriately over 70% of the time. However, items such as diagnosis, patient education, and procedures were undocumented up to 50% of the time. A study that compared the written record with a videotape recording of patient encounters found that only 56% of medical history items were recorded in the chart. Psychosocial issues were even less likely to be recorded.\textsuperscript{41} Leaf and colleagues found that chart audit correlated poorly with physician self-report in a study of cardiac risk assessment in an outpatient clinic.\textsuperscript{42}

Two more recent studies have highlighted the potential problems using medical record audits to measure specific aspects of quality of care and performance. One study compared medical record audit with a standardized patient’s (SP) completion of a checklist for quality-of-care items. Luck and associates found that the overall quality score from an SP was significantly higher than the chart audit. In this study, the medical record audit was only 70% specific when compared with the SP as the gold standard.\textsuperscript{43} The same authors then compared clinical vignettes, SPs’ scores, and medical record audits and again found that medical record audits detected the lowest number of quality criteria among a group of faculty and residents.\textsuperscript{44} A study by Ognibene and colleagues\textsuperscript{45} is one of the few to find a high rate of documentation: 81% of residents successfully documented 10 components of the physical
examination in the medical record. At the present, the majority of studies using written records demonstrate that many important aspects of the medical encounter are not recorded in the medical record. Electronic medical records may improve documentation of items such as physical examination but this has yet to be proved, and nontechical aspects of care, such as counseling, may still suffer from poor documentation.

A “good chart” does not necessarily equal “good care.” For example, the chart may have a check box for smoking cessation counseling, but such a “check” does not provide much information about what was covered in the counseling session. More work is needed on examining the impact of quality charting with patient outcome. This may be especially important in an era when patient care is often fragmented among a number of doctors who “communicate” diagnostic and therapeutic choices through written records that include letters and e-mail. Lack of continuity is an especially pressing problem for residency training programs. These studies raise questions as to the best combination of methods to measure both trainees and program performance regarding quality of care.

Process versus Outcomes

Medical records are a reasonably good method to determine if specific processes of care have been performed, especially when explicit criteria are defined. However, the utility of using the medical record audit to determine causation for patient outcomes is very limited. Most often a surrogate outcome is used such as blood pressure, hemoglobin A1c level, absence of a postsurgical complication, etc. Systematic approaches to reviewing critical incidents, such as root cause analysis, will use information from the medical record.

Implicit Review

Reviewing a medical record without a minimal framework, structure, or especially well defined criteria results in low reliability and reduced validity. Much of chart review in quality improvement programs previously utilized implicit review by “experts.” The reliability of implicit review, also known as peer review, came under attack because of low reliability and the resulting negative perceptions of these reviews by physicians. Goldman reviewed 12 studies investigating quality of care and found agreement among reviewers was poor, and often only marginally better than chance. Hayward and associates examined the quality of structured implicit review in evaluating the care of inpatients on a general medicine service. Reasonable agreement (kappa = 0.5) was seen only with overall assessment of care and prevention of death. They also noted that at least five reviewers per patient were necessary to achieve an accuracy of 90%. This poor inter-rater agreement was noted despite 15 hours of training for the select group of physician reviewers! They concluded implicit review, even with structured criteria linked to a rating scale, was “expensive, burdensome, and untenable for many specific quality-of-care judgments.” Because of these observations and others, the CMS abandoned the peer review approach to measuring quality of care.

Although many faculty members probably do recognize a “poor chart” when they see one (the gestalt factor), the lack of defined criteria as to why the chart is poor is a disservice to the trainee, who needs specific feedback in order to improve. Furthermore, numerous studies in other settings have found that implicit reviews (e.g., those without predefined, explicit criteria for abstraction) contain too many errors. Thus, simply reviewing a chart for “quality” without predefined criteria and objectives will have low reliability and is much less likely to provide useful information about performance or for feedback to the trainee.

Assessment of Clinical Judgment

Resident analytic and integrative skills can be assessed only partially through record review, especially when one considers the problems in the quality of documentation. Furthermore, is the physician’s judgment adequately recorded on the record? Did that judgment translate into an appropriate management plan? Gennis and Gennis found that when a separate attending physician evaluated a patient independently from the resident, the faculty attending’s recommendations for care management was different in nearly 33% of the resident’s patients. A similar study in a military outpatient training clinic found a similar frequency of differences between attending and resident management decisions but the differences were less dramatic and the majority of the recommended changes from faculty were minor. These two studies raise significant questions about the ability to accurately assess the appropriateness of management plans from medical record review. Chart-stimulated recall, discussed later, is a method to more effectively use the medical record to assess clinical judgment.

Time

Medical record review can be very time-consuming, especially if it is used for high-stakes decisions. Norman and Ognibene both found that audit of a large number of charts is needed to ensure reasonable reliability in the training setting. Researchers working with practicing physicians have found that at least 25 to 35 patient medical records are required for pass-fail decisions for a single condition (e.g., diabetes) in provider recognition programs. Audits require the development and testing of abstraction instruments, data collection and entry, data analysis, and then dissemination of the results to individual residents. However, several factors can help to minimize these limitations. First, I strongly recommend using standardized abstraction tools and quality measures already developed and field-tested whenever possible. In the United States, there is actually a Web-based clearinghouse for quality measures supported by
the Agency for Healthcare Research and Quality (AHRQ). The National Quality Forum systematically endorses quality measures; approved measures can be downloaded from their website. An example of an abstraction form for preventive services and diabetes care developed and field tested by Qualidigm, a QIO in the United States, is provided in Appendix 5-1. Using existing well-defined measures and abstraction tools can save training programs substantial time. Second, considering having the trainees perform the actual audit. Not only does this save time for faculty and programs, but as previously discussed, the self-audit experience is valuable for the trainee.

**Cost**

Cost may be a factor in your program if the medical record audit is performed on archived records and the institution charges a fee for pulling the charts. Cost will also be a factor if you use faculty or other administrative personnel to perform the abstraction. For faculty, the usual cost is their time. If you use abstractors, they may charge a monetary fee for their services.

**Faculty Development**

We have found few current faculty members who have extensive experience with medical record audits. There are several key issues around faculty development: personal skill documentation, abstraction skills, and interpretation. Many faculty members exhibit the same behaviors as trainees when documenting the results of their own medical encounters. Furthermore, some faculty members are learning to use electronic medical records at the same time as the trainees. When faculty skills in using medical records are suspected to be problematic your first priority should be to train your faculty in the optimal use of the medical record at your institution. Second, reliable and accurate abstraction is a skill in itself, and most faculty members have little experience. While I do not advocate that faculty be the primary source for abstraction services, faculty members do need to understand how to conduct a proper audit, including how to use an abstraction manual and how to properly interpret the specifications of quality measures. Finally, the faculty needs to know how to interpret the results of an audit to help trainees improve. For example, what should a faculty member tell a trainee whose “quality report” shows poor compliance with several quality measures? Chapter 11 on practice-based learning and improvement provides guidance and suggestions. Brook and Lohr appropriately pointed out that when using chart review as an educational tool you must have committed faculty and the specific goals with explicit criteria. This is essential in any review process.

3. Ultimately, collection of data directly from the patient is important, especially when considering “quality of care.” This last point reinforces the need for a multifaceted approach to evaluation, highlighting how combining direct observation and patient surveys with medical record audit can be a potentially powerful combination for assessment.

**Summary of Limitations of Medical Record Audits**

1. Medical record review can be time-consuming, and to be reliable for “high-stakes” decisions, review of a substantial number of records (usually more than 25 medical records) per trainee is needed. A substantial commitment from faculty may be needed depending on local information systems and resources. As electronic medical records evolve with better search capabilities and registry functions, real-time audits may be able to provide continuous quality feedback.

2. The quality and completeness of documentation hamper validity of medical record review. The written record rarely records the physician-patient interaction comprehensively.

3. Medical record audits cannot assess the quality of important aspects of the encounter. For example, documentation of the cardiac examination says nothing about the skill of the examiner. Quality of patient instruction by the physician cannot be assessed.

4. It is difficult to adequately assess physician interpretive abilities and judgment without corroboration.

5. The current lack of consistency in medical record organizational format will continue to hamper the standardization of medical record audits.

6. Medical record review can be potentially costly if the audit is performed by trained individuals outside the training program.

7. Medical records may be better for assessing the process of care and be less useful when looking at patient outcomes. Specific measures (e.g., compliance with preventive health measures such as immunization) can help enhance the utility of chart review.

8. Implicit review, even if “structured,” suffers from significant reliability problems. Furthermore, the investment required for training is significant. Table 5-1 summarizes some of the key limitations of medical record audits with possible solutions.

**Chart-Stimulated Recall**

How can medical records be used to more effectively assess clinical judgment? Chart-stimulated recall (CSR)
uses either a single medical note or multiple components of the medical record to assess a trainee’s analytic, integrative, and clinical judgment skills. CSR enables a faculty rater to assess a trainee’s rationale for diagnostic and treatment decisions, the other options that were considered but ultimately discarded, and the reasons these other options were ruled out. CSR also allows the rater to investigate any other factors that influenced the trainee’s clinical decision making, such as patient, environmental, and microsystem factors (see Chapter 11). The medical note, including admission, acute encounter, daily progress, and routine outpatient follow-up notes, serves as a template for the examiner (e.g., faculty) to develop specific interview questions that probe a trainee’s clinical decision making and judgment skills.

The American Board of Emergency Medicine (ABEM) performed pioneering work with this technique in the 1980s. The ABEM used CSR as a way of constructing an oral examination using emergency department records from the examinee.20 The ABEM examiners used the examinee’s charts to assess knowledge, clinical problem solving, and a “general measure of clinical competence.” The ABEM ultimately found that the CSR was reliable and valid.21 In fact, the ABEM found that three to six charts could provide sufficient reliability for CSR interviews. However, from a certification, high-stakes testing standpoint the CSR was too expensive and time-consuming. Despite their satisfaction with the approach, the ABEM discontinued CSR as part of its certification process.20,21

Jennett and Affleck provided a number of important recommendations and guidelines for performing CSR. First, the CSR interviewer should ideally be a medical faculty member and be trained to perform CSR consistently and reliably.22 We recommend using the techniques and principles of performance dimension and frame of reference training described in discussion of direct observation (Chapter 9). Second, the medical record of interest should be reviewed beforehand to develop the specific questions for the CSR interview. Third, the interview should begin with open-ended questions to minimize rater bias and to provide the rater with insight into the trainee’s thought processes. All questions should be asked in a nonjudgmental, nonbiased approach. Depending on the intended use of the CSR, the session can be audiotaped for future analysis. Audiotaping also allows the rater to focus more on the interview and questions instead of taking the time to document the discussion and decisions. CSR can be performed in person or by telephone, if necessary.

At Yale University, we used CSR for formative assessment and as part of our diagnostic approach for residents in difficulty (see Chapter 13). In our CSR approach, we used single encounter notes (admissions, daily progress notes, outpatient notes, etc.) for two main purposes. First, we look at the basic “quality” of the notes with a simple framework. Questions for the reviewer were: Is the note legible? Does it follow a standard format (SOAP or problem-based)? Second, as part of the “quality” review, we examine notes for two specific deficiencies: lack of internal consistency and discordances. Internal consistency refers to whether an issue or problem noted in the history or physical examination section of the note is logically “carried through” the remainder of the note (see Appendix 5-2). For example, if the trainee lists and describes in the subjective or history section of the note chest pain as a symptom, then the physical examination, assessment, and plan sections should all contain pertinent information about that symptom. The physical examination should contain the appropriate components relevant to this symptom (e.g., cardiac examination) and the assessment and plan should provide a differential diagnosis and plan of action for the chest pain, respectively.

Discordance refers to a phenomenon that occurs when information in one section of the medical note is discordant with information or decisions documented.

| Table 5-1 Summary of Medical Record Audit (MRA) Limitations |
|-----------------|-----------------------------------------------------|
| **Limitation** | **Possible Solutions** |
| Quality of documentation | Use problem lists and flowcharts for chronic conditions and preventive care. Provide templates for medical history and physical examination. Use electronic medical record (may or may not improve documentation; training required in effective use of EMRs). Combine MRA with direct observation. |
| Time | Have trainees perform audit of their own charts or their peers’ charts. Seek assistance from hospital or clinic quality improvement department to generate performance reports, especially if have EMR. Use other health care personnel (if available). |
| Implicit review | Provide minimal framework for medical record review and do not rely solely on the judgment of the reviewer. Encourage explicit criteria whenever possible. Provide auditor training. |
| Cost | Have trainees perform audit of their own charts or their peers’ charts. Use existing reports, when available, from quality improvement departments. |
| Assessing clinical judgment | Combine MRA with chart-stimulated recall. |
in other sections of the note. Returning to our chest pain example, if a trainee lists chest pain with accompanying signs and symptoms that suggest evidence of ischemic disease but lists heartburn as the likely diagnosis without mention of coronary artery disease, the assessment is “discordant” with the history information provided.

Although problems with internal consistency and discordance may simply be related to documentation errors, deficits in knowledge and clinical decision making are more likely. The faculty member performing CSR can use these observations to probe a trainee’s medical knowledge and clinical decision-making skills. The roster of questions developed for the Canadian physician assessment review (PAR) CSR is an excellent template to help guide faculty. CSR can still be a useful technique even if a chart doesn’t contain problems with internal consistency or a discordance. For example, a trainee may have appropriately diagnosed hypertension and chosen a “diuretic” for treatment. Choice of a diuretic would certainly be an evidenced-based choice. However, was the decision to prescribe a diuretic simply a “rote” choice or was the trainee aware of the related guidelines? Was a diuretic the best choice for this patient? Is the trainee aware of the side effect profile of the medication and how it may relate to the specific patient’s risk factors and other comorbidities?

CSR can be “combined” with other evaluation tools to measure competence. For example, faculty members can combine a direct observation exercise, such as the mini-clinical evaluation exercise, with a CSR exercise. Information obtained from direct observation can be combined with the medical record documentation to assess deeper aspects of clinical decision making. Combining CSR with medical record audit can provide a very robust assessment of the quality of care at the individual patient level. A study in Canada found that a combined medical record audit and CSR interview for a small group of practicing family practitioners was a valuable experience for assessing and improving quality of care.

Also, in a number of continuing education settings CSR has been found to be a useful tool for both assessing the impact of a continuing medical education course and determining learning needs for practitioners.

**Conclusion**

Medical record audits and data feedback can be valuable tools to assess clinical competence. Given the critical importance of performance data for quality improvement and the competency of practice-based learning and improvement, all trainees should receive individual performance data at a minimum during training. Medical records are also readily accessible. They allow the examination of a potentially large number of clinical encounters, and their use in assessment may be relatively unobtrusive, thereby minimizing the Hawthorne effect. When explicit criteria and end points are used, such assessments may yield important information about practice habits in specific areas of care (e.g., preventive health). Databases and EMR may provide a wealth of accessible information in a timely and ongoing manner. Furthermore, the Residency Review Committee requires medical record reviews. Use of the chart-stimulated recall technique helps educators to evaluate more complex skills such as clinical judgment. Given the new structure of the ACGME general competencies, more effective use of the medical record, whether paper or electronic, will be needed for program directors.

Involving trainees in medical record audit activities is strongly recommended. As the level of scrutiny for physicians increases through health care insurance organizations, Center for Medicare and Medicaid Services (CMS), and peer review organizations, to name just a few, physicians will need a good understanding of audit methodology, and chart review remains a cornerstone of most audit programs. Therefore, this skill should be incorporated into residency training. The quality of care audit provides a simple example of a program that can be incorporated into a residency program. This is also an excellent method to satisfy the requirements of the new practiced-based learning and improvement competency while simultaneously working to improve outcomes among patients cared for by the house staff program.

**ANNOTATED BIBLIOGRAPHY**


   This comprehensive review of the effects of medical record and feedback on physician behavior includes a valuable summary of studies from training settings. The authors used the rigorous methodology from the Best Evidence Medicine Education (BEME) initiative (see http://www.bemecollaboration.org). Consistent with previous studies, medical record audit and feedback can produce modest changes in quality.


   Little is known about the impact and value of trainees performing their own audits instead of simply receiving data from an external source. Both preceding studies found the process of self-audit was a powerful experience for residents and practicing physicians. Both groups experienced a “surprise” at the gap between knowing what to do and their actual clinical performance. Both studies provide guidance on how to set up self-audit. The abstraction instrument used in the resident study (article 2) is provided in Appendix 5-1.


   This remains one of the best reviews of all the work done on chart-stimulated recall. This is a valuable article for any educator planning to implement chart-stimulated recall in their program.


   The PAR program in Alberta has developed a useful, practical form for conducting chart-stimulated recall. The form comes with clear instructions and provides a series of possible questions the reviewer can ask in a chart-stimulated recall. This form has been field tested and studied by investigators working with the College of Physicians and Surgeons of Alberta. I’d recommend educators download the pdf CSR form if they are considering implementing CSR in their program.
REFERENCES


APPENDIX 5-1

Sample Medical Record Abstraction Form for Diabetes

| Date of abstraction | __________________________ | Physician name | __________________________ |
| Abstractor initials | __________________________ | First visit with this MD? | YES | NO |
| Abstraction time (min) | __________________________ | First visit to this clinic? | YES | NO |

Demographics

| 1. Patient Name (First, MI, Last) | __________________________ |
| 2. Patient Identification (MRN) # | __________________________ |
| 3. Gender (circle) | a. Male | b. Female |
| 5. Date of birth | __________________________ (MM/DD/YYYY) |

Chart Information

| 6. Does the chart contain a problem list? | Yes | No |
| 7. Does the chart contain a preventive services checklist? | Yes | No |
| 8. Does the preventive services checklist have any entries? | Yes | No |

Conditions Present

| 9. Hypertension | Yes | Not documented |
| 10. Coronary artery disease | Yes | Not documented |
| 11. Heart failure | Yes | Not documented |
| 12. Conduction disorder/bradyarrhythmia | Yes | Not documented |
| 13. Aortic stenosis | Yes | Not documented |
| 14. Chronic obstructive pulmonary disease/asthma | Yes | Not documented |
| 15. Cerebrovascular disease | Yes | Not documented |
| 16. Peripheral vascular disease | Yes | Not documented |
| 17. Chronic renal disease | Yes | Not documented |
| 18. Chronic liver disease | Yes | Not documented |
| 19. Diabetes mellitus | Yes | Not documented |
20. Dyslipidemia Yes Not documented
21. Breast cancer Yes Not documented
22. Colon cancer Yes Not documented
23. Dementia Yes Not documented
24. Bleeding disorder/risk Yes Not documented
25. Peptic ulcer disease Yes Not documented
26. Anemia Yes Not documented

**Physical Examination**

27. Height ________ Inches Not documented
28. Date of most recent height _ _ / _ _/ _ _ _ _ N/A
29. Weight ________ lbs Not documented
30. Date of most recent weight _ _ / _ _/ _ _ _ _ N/A

**Physical Examination**

31. Record all blood pressures from the last three visits during the observation period (insert period).

<table>
<thead>
<tr>
<th>Date</th>
<th>Blood Pressure</th>
<th>Date</th>
<th>Blood Pressure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. _ _ / _ _</td>
<td><strong><strong><strong>/</strong></strong></strong>_</td>
<td>7. _ _ / _ _</td>
<td><strong><strong><strong>/</strong></strong></strong>_</td>
</tr>
<tr>
<td>2. _ _ / _ _</td>
<td><strong><strong><strong>/</strong></strong></strong>_</td>
<td>8. _ _ / _ _</td>
<td><strong><strong><strong>/</strong></strong></strong>_</td>
</tr>
<tr>
<td>3. _ _ / _ _</td>
<td><strong><strong><strong>/</strong></strong></strong>_</td>
<td>9. _ _ / _ _</td>
<td><strong><strong><strong>/</strong></strong></strong>_</td>
</tr>
<tr>
<td>4. _ _ / _ _</td>
<td><strong><strong><strong>/</strong></strong></strong>_</td>
<td>10. _ _ / _ _</td>
<td><strong><strong><strong>/</strong></strong></strong>_</td>
</tr>
<tr>
<td>5. _ _ / _ _</td>
<td><strong><strong><strong>/</strong></strong></strong>_</td>
<td>11. _ _ / _ _</td>
<td><strong><strong><strong>/</strong></strong></strong>_</td>
</tr>
<tr>
<td>6. _ _ / _ _</td>
<td><strong><strong><strong>/</strong></strong></strong>_</td>
<td>12. _ _ / _ _</td>
<td><strong><strong><strong>/</strong></strong></strong>_</td>
</tr>
</tbody>
</table>

**Counseling/Prevention**

32. Was an assessment of tobacco use performed? Yes Not documented
33. Is patient a current smoker? Yes No Not documented
33a. Was smoking cessation counseling offered? Yes Not documented N/A

**Counseling/Prevention**

34. Was a foot exam performed? (insert observation period) Yes Not documented
34a. Was a monofilament test for neuropathy performed? Yes Not documented
35. Has the patient ever received pneumovax? Yes Not documented

**Labs/Diagnostic Studies:**

For all questions pertaining to labs/diagnostic studies, review the record from 6/30/2001 back to 7/1/1999 (if necessary). Record the most recent date that the test was performed prior to 6/30/2001 and the value.

<table>
<thead>
<tr>
<th>Lab Study</th>
<th>Test/Diagnostic Study</th>
<th>Test Performed</th>
<th>a. Date Performed</th>
<th>b. Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>36. Blood urea nitrogen</td>
<td>Yes Not documented if yes, record the date and value.</td>
<td>36a. _ _ / _ _ _ _</td>
<td>36b. ___________ (Normal range 6–19 mg/dL)</td>
<td></td>
</tr>
<tr>
<td>37. Creatinine</td>
<td>Yes Not Documented if yes, record the date and value.</td>
<td>37a. _ _ / _ _ _ _</td>
<td>37b. ___________ (Normal range 0.6–1.4 mg/dL)</td>
<td></td>
</tr>
<tr>
<td>38. Blood sugar</td>
<td>Yes Not documented if yes, record the date and value.</td>
<td>38a. _ _ / _ _ _ _</td>
<td>38b. ___________ (Normal range 70–105 mg/dL)</td>
<td></td>
</tr>
<tr>
<td>39. Blood sugar recorded as fasting?</td>
<td>Yes Not documented</td>
<td>39a. _ _ / _ _ _ _</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Labs/Diagnostic Studies—cont’d

<table>
<thead>
<tr>
<th>Lab Study</th>
<th>Test/Diagnostic Study</th>
<th>Test Performed</th>
<th>a. Date Performed</th>
<th>b. Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>40.</td>
<td>Albuminuria test</td>
<td>40. Yes Not documented</td>
<td>40a. <em>/</em> _/ _/ _/ _/</td>
<td>40b. Albumin present? Yes Not documented</td>
</tr>
<tr>
<td>41.</td>
<td>Hemoglobin A\textsubscript{1c}</td>
<td>41. Yes Not documented</td>
<td>41a. _/ _/ _/ _/ _/</td>
<td>(Normal range 3.0–6.5%)</td>
</tr>
<tr>
<td>42.</td>
<td>Total cholesterol</td>
<td>42. Yes Not documented</td>
<td>42a. _/ _/ _/ _/ _/</td>
<td>(Normal range 120–220 mg/dL)</td>
</tr>
<tr>
<td>43.</td>
<td>HDL cholesterol</td>
<td>43. Yes Not documented</td>
<td>43a. _/ _/ _/ _/ _/</td>
<td>(Normal range 44–55 mg/dL)</td>
</tr>
<tr>
<td>44.</td>
<td>LDL cholesterol</td>
<td>44. Yes Not documented</td>
<td>44a. _/ _/ _/ _/ _/</td>
<td>(Normal range 40–170 mg/dL)</td>
</tr>
<tr>
<td>45.</td>
<td>Triglycerides</td>
<td>45. Yes Not documented</td>
<td>45a. _/ _/ _/ _/ _/</td>
<td>(Normal range 40–150 mg/dL)</td>
</tr>
<tr>
<td>46.</td>
<td>Potassium</td>
<td>46. Yes Not documented</td>
<td>46a. _/ _/ _/ _/ _/</td>
<td>(Normal range 3.3–5.1 mEq/L)</td>
</tr>
<tr>
<td>47.</td>
<td>EKG performed</td>
<td>47. Yes Not documented</td>
<td>47a. _/ _/ _/ _/ _/</td>
<td>(Select all recorded findings)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>a. Myocardial infarction (any age)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>b. Atrial fibrillation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>c. LVH</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>d. LBBB</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>e. None of the above</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>f. No interpretation</td>
</tr>
</tbody>
</table>

Treatment

48. Does the chart contain a current list of medications? Yes Not documented

48a. Are the patient’s medications documented at the last visit to this physician? Yes Not documented

49. Record all medications that the patient was taking or that were prescribed at the end of the observation period (insert observation period). Use hospital discharge summaries, consultation notes, phone conversations, etc., if necessary.

50. Does the chart contain a medication allergy section? Yes Not documented
Office Visits (Measurement Year):

51. Record all dates on which the patient was seen at this office (Chase) during the observation period.

1. __/__/____
2. __/__/____
3. __/__/____
4. __/__/____
5. __/__/____
6. __/__/____
7. __/__/____
8. __/__/____
9. __/__/____
10. __/__/____
11. __/__/____
12. __/__/____
13. __/__/____
14. __/__/____
15. __/__/____
16. __/__/____
17. __/__/____
18. __/__/____
19. __/__/____
20. __/__/____
21. __/__/____
22. __/__/____
23. __/__/____
24. __/__/____
25. __/__/____
26. __/__/____
27. __/__/____
28. __/__/____
29. __/__/____
30. __/__/____

Created by Qualidigm and the Yale Primary Care Internal Medicine Residency Training Program.

APPENDIX 5-2

Basic CSR Documentation Template

Chart-Stimulated Recall Note

Trainee: ____________________________ Date: ____________________________

Level

1. Organization and clarity of note
   (Is the note organized? Is the format appropriate and consistent? Is it legible?)

2. Note content
   (Are clinical issues explained in sufficient detail? Is any essential information missing?)

3. Internal consistency
   (Do clinical issues follow a logical sequence throughout the note?)

4. Discordances
   (Are there any maneuvers or decisions that are discordant with other information provided in the note?)

5. Questions for resident
   (Write up to five questions you would want to ask this resident with a focus on clinical reasoning and judgment.)
   A.
   B.
   C.
   D.
   E.