Using Utilization Measures to Improve Quality in Health Systems

Stephen M. Davidson, Ph.D.
Professor
Boston University School of Management
595 Commonwealth Avenue
Boston, MA 02215
sdavidso@bu.edu 617-353-7422 617-353-5244 (fax)

Prepared for the Center for Health Management Research, October 2005

Using Utilization Measures to Improve Quality in Health Systems

Stephen M. Davidson, Ph.D.

It is widely acknowledged that the safety and quality of health care in the U. S. is not as high or consistent as it should be. As a result, health system leaders are being called on to improve the quality of care provided in their systems. Before effective strategies can be devised and implemented to improve conditions, however, they need to know exactly what the state of quality is in their systems. Only then can institutional leaders act effectively with knowledge to set specific goals they should attempt to reach and be able to assess the degree to which they have accomplished those goals. By "the state of quality" we mean (1) the extent to which care provided in a particular system meets established criteria for good quality care and (2) specific deficits that can be identified and act as targets for improvement.

This paper is organized into two sections, one for each of its two main purposes. The first is to examine factors that could affect the suitability of using utilization measures in efforts to measure quality of care delivered in their systems as a necessary step in the effort to improve safety and quality of care in health care organizations or systems. The second is to discuss ways of actually using those measures to improve quality of care in those systems.

First quality will be defined so that it can be measured and reflect the care provided in the health care system. Utilization measures will be related to quality and allow better quality to be distinguished from lesser quality.

There are two principal sources of utilization measures: claims for services rendered and medical records. The advantages and limitations of both will be discussed. Additional data reflecting demographic characteristics of patients as well as diagnosis and other indicators of health status before and after treatment will also be discussed. Among the uses to which these data can be put are the following: (1) determine whether or not indicated services were provided; (2) identify patterns of services provided and compare to condition-specific standards of care; and (3) determine whether indicated follow-up services were provided and when.

Having produced utilization measures and interpreted them to represent the quality of care provided, we will turn to ways in which the data can be used by clinical and other leaders of health systems to improve safety and quality of care.

Using Utilization Measures to Improve Quality in Health Systems

Stephen M. Davidson, Ph.D.
Boston University School of Management
October, 2005

1. A Problem with Quality of Care in the U. S.

It is widely acknowledged that the safety and quality of health care in the U. S. is not as high or consistent as it should be. In a chartbook prepared for the Commownealth Fund, Leatherman and McCarthy present evidence of indicators related to a number of key issues, including effectiveness and patient safety.[1] For example, they show large numbers of children who were not up-to-date on five of the recommended immunizations,[2] Medicare beneficiaries hospitalized for pneumonia who had not received timely and appropriate antibiotic care,[3] and patients who had not received recommended care for their diabetes.[4] In addition, citing the Institute of Medicine, they report that between 4,000 and 98,000 people died in 1997 from medical mistakes as well as disconcertingly large numbers of patients who suffered from preventable adverse drug events.[5]

HEDIS results show that, although progress has been made at reducing quality deficits in recent years, large problems still exist. For example, NCQA reports comparisons between the national averages among health plans on a variety of measures with the results for the top 10 percent of plans that demonstrate "enormous quality gaps remain."[6] Among the largest are controlling high blood pressure and cholesterol management. In both cases, a spread of more than 20 percentage points was observed between the national average among all plans and those for plans at the 90th percentile.[7] In fact, these figures understate the problem since it must be assumed that many plans fall well below the national average among plans.

Because of reports like these, many health systems, providers of services, payers for care, and others are interested in trying to improve the level of safety and quality. Indeed, many institutions have already taken steps to improve the quality and consistency of their services, and it is reasonable to assume that others will do so, as well.

Before clinical and other leaders can devise and implement effective strategies to accomplish that goal for their organizations, however, they need to know the state of quality in their systems. By "the state of quality" we mean (1) the extent to which care provided in a particular system meets established criteria for good quality care and (2) the site of specific deficits. Only then can institutional leaders act effectively to set specific goals for their own systems and then, later, to assess the degree to which they have accomplished those goals.

This paper has two main purposes and is organized into two sections, one for each. The first purpose is to examine factors that could affect the suitability of using utilization measures to improve safety and quality of care in health care organizations or systems. Then, having accomplished the first purpose, the second is to discuss ways of

actually using those measures to improve quality of care. Therefore, in Part One, first, we will define quality and its measurement, with particular emphasis on utilization measures. Then, we will assess issues associated with the two main sources of utilization measures, medical records and claims for services. Next, given the limitations of utilization measures, we will ask what can reasonably be expected from them in the effort to measure quality of care. Finally, we conclude with some summary observations. Then, in Part Two, having identified and resolved the challenges regarding the use of utilization measures and having assessed the state of quality in their organizations, we discuss ways in which health plan leaders can use those utilization measures as tools in a concerted effort to improve quality within those same health care organizations.

PART ONE: UTILIZATION MEASURES

2. What is Quality in Health Care?

Quality is a complex concept. It encompasses the extent to which appropriate services are used, the skill with which services that are used are provided, and the relationship of those services to achieving desired clinical outcomes. Although quantifying it with any degree of precision is difficult,[8 9] its many aspects can be indicated with a variety of measures.[10 11 12] In this section, we will identify and define it and discuss how to measure it.

Quality has been defined as " 'the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge.' Good quality means providing patients with appropriate services in a technically competent manner, with good communication, shared decision making, and cultural sensitivity."[13] In other words, given the state of medical knowledge and the patient's condition, good quality care consists of providing those services most likely to improve the patient's health. The definition includes the selection of services to respond to a patient's condition, the skill with which those services are delivered, the context in which they are provided, and the outcomes they produce. It implies that the ultimate criterion is the impact of services on health outcomes, but by noting that the test is not the presence of desired outcomes for each individual, but "the likelihood of desired health outcomes," it also implies that other factors besides health services determine a person's health status. Partly as a result, the studies needed to assess the quality of care must compare *groups* of patients who have a given diagnosis and who receive care with certain characteristics to groups of similar patients whose care does not have those characteristics. Then, the test of quality is whether the probability of good outcomes is higher for the first group than the second.

How does one know when the care provided is of good quality? It is not something one can recognize just by looking at a doctor or a clinical team providing services. Rather, it depends on whether the services provided by those people produce a desired result – the patient with pneumonia gets better, the patient with a broken leg can walk again, or the patient with chronic pain is more comfortable. Moreover, for any one patient, the result may be months or even years in the future – for example, a patient still

cancer-free five years post-treatment or a diabetes patient whose eyesight has not deteriorated with passing years. Measurement implies counting something, but given the situation just described, what do we count? And how can we be sure that what we are counting, first, has any relation to quality at all and, second, distinguishes between good quality and lesser quality?

The classic conceptualization of the challenge was articulated by Donabedian,[14] who argued that quality of care could be assessed on the basis of three characteristics: its structure, process, and outcomes. He wrote that the probability of good care increases when certain *structural conditions* are in place, including appropriate facilities and equipment and well-trained professionals. Patients obtaining care from good facilities provided by clinicians known to be competent are likely to be better served than similar patients whose care was provided elsewhere. These structural characteristics do not guarantee good quality, but they increase the chance that it will occur. Similarly, if the *processes of care* meet certain standards, the chances increase that the care obtained is of good quality, as well. In other words, if a history is taken, tests appropriate to the patient's symptoms are performed, and prescriptions or procedures that are consistent with the resulting diagnosis are provided, the probability of a desired outcome improves.

While structure and process are important, the ultimate test is the actual impact of the services received on patient health status (that is, on *clinical outcomes*). In other words, following the delivery of appropriate services, if a patient's health status does improve or is consistent with the desired condition-specific outcome and if the improvement can be attributed to the services received, the ultimate test will have been met. Implicitly in Donabedian's conceptualization, however, if one cannot measure outcomes directly or soon or cannot attribute them to the services provided, the next best thing to do is to measure whether or not services determined in prior studies to be appropriate were actually provided (that is, *process*). And if it is not possible to know exactly which services were provided (that is, what process really was used), then, indicators of *structure* are better than nothing. The bottom line is that measuring quality is difficult, and compromises must often be made.

While the basis for assessing the effects of care on health is clear conceptually, in practice it is difficult to demonstrate that measurable changes in health status are actually attributable to medical care received for at least two reasons. One is that many conditions are self-limiting, and people tend to get better eventually even without care. So, did a particular person's health improve *because* of the care he got? Did good care allow him to recover sooner than he would have otherwise, and if so, how much sooner and with what other benefits and costs? Or was the care irrelevant because he would have got better anyway? Determining the effects of care on people with chronic conditions is even harder because they tend not to get "better." Rather, for them, success is an adequate or improved level of functioning, the stability of their condition over time, the slowing of the pace of inevitable deterioration, or the absence of pain.

The second reason it is hard to demonstrate the effects of medical care is that since care is only one element affecting improved health status, the determination of quality can be done only with studies that are designed carefully to reveal actual

differences in outcomes between similar *groups* of patients that received two different levels of care. In other words, for a particular diagnosis, did a higher proportion of patients in the group who obtained a specific set of services show improvement than in the group who had different care or none at all? To do that requires a strong study design, which includes: (1) selecting patients whose baseline conditions are similar (e.g., have the same diagnosis), (2) dividing those patients into experimental and control groups whose primary difference is in the care received, (3) ensuring that all patients in the experimental group received care that met the criteria being tested and all patients in the control group did not, (4) allowing enough time to pass for the care to have the anticipated effect, (5) ensuring that the two groups are large enough that even small differences in the probability of different outcomes are actually revealed, and (6) having appropriate data regarding the structure or processes of care, and/or outcomes to determine the effects of care.

And this brings us back to measurement. When we design these studies, what data do we want and where can we get it? Since the underlying research strategy is to determine the extent to which people ill with a given condition who get certain services are more likely to improve than similar people who do not receive those services (that is, who use different services or none at all), we need utilization data.[*]

There are two principal sources of utilization measures: claims for services rendered and medical records. Both sources should provide, first, evidence as to which services were used, if any, as well as descriptive information about the patients whose utilization is being studied – that is, demographic characteristics as well as diagnosis and other indicators of health status before and after treatment. We turn next to a discussion of utilization data.

3. Issues with Utilization Data

The first point to make about both sources of utilization data is that they were created for other purposes and not to determine the quality of care by measuring its effects on groups of patients. *Claims* exist primarily so that the provider of services – whether an individual clinical professional or a health care organization – can be paid for work performed. *Medical records*, on the other hand, were created primarily so that the patient's doctor (usually, in the singular) will have a running record of the patient's history, condition, and services provided, primarily to help that same doctor and others who may care for him or her in his future work with the patient. While each may achieve its primary goal admirably, five factors determine its adequacy for assessing the quality of care. Two of those items relate to the content of the data record – the completeness and accuracy of the information in the source document; and three items affect the burden that must be shouldered for the information to be useful in assessing quality of care – the

[*] We need data that allows us to know the two groups are similar in their health status at baseline. Either they need to have the same diagnosed illness or, if we want to test the effects of preventive services, they need to be equally healthy. Then, we need data to show what services they received and what their health status was following receipt of the services. The dependent or outcome variable will be the change in their health status; and the independent variable will be the care they received.

retrievability and confidentiality of the data and its suitability for analysis. We will discuss each in turn. They are summarized in the following table.

Factor	Medical Records	Claims
CONTENT		
Completeness	Detailed description or sketchy notes? Services provided by others? Similar data for all patients?	Unreimbursed services? Services covered by the deductible? Services provided by others? Similar data for all patients?
Accuracy	Defensive medicine? Reliability of patient-entered data and clinician judgments or opinions?	Upcoding?
USEABILITY BURDEN		
Retrievability	Electronically by computer? Manually from a file room?	Electronically?
Confidentiality	Remove identifying information when transferring data.	Remove identifying information when creating files for analysis.
Suitability for Analysis	Data extraction by hand. Enter onto forms and/or computer.	Program algorithms to select suitable cases and extract relevant data.

Table: Factors Affecting the Usefulness of Utilization Data Sources for Measuring Safety and Quality

Completeness – The fundamental question is, do the claims or medical record include the patient's entire utilization experience? Theoretically, at least, any items that are omitted may explain the difference in outcomes attained by the two groups, especially if one group had some which the other group did not. Having said that, there are reasons to be concerned about both sources. Since claims exist primarily to result in a payment from an insurer or other third party, they may include only reimbursable services. They may omit valuable information about services used to satisfy a deductible, if any, and services provided by other clinicians. On the other hand, since medical records are created by and mainly for the benefit of the patient's physician, they are likely to be idiosyncratic. The entries of some physicians (and others who input data) may be comprehensive, including detailed notes, while those of others include only snippets of information to jog their memory. Moreover, often medical records are not intended for use by others. Even when a physician's partner covers on a day off or during a vacation, often the clinical goal is simply to maintain some degree of normalcy until the primary physician returns. Further, specialists whose patients share a narrow range of clinical conditions are likely to have different records than generalists whose patients cover a much broader range of ailments. Medical records follow no standard format, and if the

patients in the study group are treated by multiple doctors, it is likely that more information will be available for some patients than for others. One result is that some patients in the original representative sample may need to be excluded from the study because certain needed data are missing from their medical records. Another issue regarding medical records is how soon after treating the patient does the clinician enter information into the record. Many physicians jot down notes during the day as they see patients and then, spend an hour at the end of the day making entries into the actual records. Those doctors who need to reconstruct what they had done during visits that had occurred many hours earlier will be dependent on their memories. Again, since multiple doctors will be involved in treating patients in each group, the original equivalence of the groups will be undermined by that fact, as well. Finally, to the extent that medical records are completed by a physician's own hand, illegibility may render parts of the records unusable and therefore, contribute to their incompleteness.

Accuracy – In addition to concerns about the completeness of claims and medical records, one must also be concerned about their accuracy. Any information source can contain random errors if people write things incorrectly, and a useful assumption is that, between groups, added or omitted services tend to cancel each other out. In the case of both of these sources, however, their other purposes create concerns about more systematic errors.

Since claims generate payments, the provider entering the data has an incentive to inflate the record of services provided. He may exaggerate the nature of some services (for example, a routine visit might be up-coded to a higher level) in order to obtain higher payment or even add services that were not actually provided. In addition to providing an historical record of a patient's condition and the services received, partly as a guide to a clinician's future actions, a medical record can also be used as evidence to show that the patient's doctor took reasonable actions according to community standards of appropriate care. In this way, he can use the record to defend himself against a claim of malpractice if the patient achieved a disappointing result. Further, what a record actually contains can be accurate, but if it is handwritten, it may be transcribed *in*accurately. This is particularly likely to occur if the handwriting of the person who made the entries is illegible. Moreover, each step in the process – from making notes to entering them into the paper record to transferring them to a computer to creating a data file for analysis – creates another opportunity for errors to occur. The fact that a given study will include records from many doctors using different processes also increases the probability that not all records in the original sample will be useable and, therefore, that the final samples will not be equivalent. An electronic record that follows a prescribed format can avoid both incompleteness and illegibility, but it is still subject to potential exaggeration or up-coding.

Confidentiality – The next limitations of the two types of records relate primarily to constraints on their usefulness that derive not from their content, but as a result of other characteristics. Assuming the accuracy and completeness of information in either the medical record or claims, neither can be used to compare groups of patients regarding quality of care without assuring the confidentiality of the person whose record is being used. If information is included that permits the individual patient to be identified, then,

the patient must give his assent to being part of the study. Those who take information from the source must remove all elements that would permit the individual to be identified by an unauthorized person for an inappropriate use. At the same time, however, the records must include descriptive information about patients that might affect their health status (e.g., diagnosis, age, gender, education level, income, family size) and, therefore, the outcome of care and that would help to characterize the composition of the experimental and control groups.

Retrievability – In order for the information in either type of utilization record to be used to assess quality of care, it must be retrievable with relative ease. Otherwise, cost will be an obstacle to its use for the purpose. Generally speaking, electronic claims and an electronic medical record are superior in this regard to handwritten, hard copy medical records. The primary reason is that the data are in predetermined fields and all the records are in the same place, a computer file. As a result, computer programs can be written to locate the sampled records and to extract from each one the relevant data regardless of whether the source is claims or medical records. Here, the key decisions concern choosing both the sample of records and the data items that need to be included.

In the case of hard copy records, on the other hand, the physical task of obtaining the records can be a challenge, first, because a method is needed to identify the full universe of appropriate records (e.g., all patients with a particular disease) and then to select a sample of records that is both representative and large enough to produce valid results. Then, once the records for the sample have been identified by number, they must be physically retrieved. If some are out of order or missing from the file cabinet where they usually reside, it may not be possible to include them. Finally, the data items need to be extracted from each record by hand, a time-consuming, labor-intensive task which will be made more difficult by the relatively free form in which information is entered.

Analysis – To be useful in the quality assessment task, the data – once obtained and assuming completeness and accuracy – must be analyzed. Again, claims and medical records in electronic form have obvious advantages. Whether from claims or medical records, data can be extracted electronically, using a common computer program, and deposited into data files for each person in the sample. Then, computer programs can be written to analyze the data and produce tables or other analytic output that tell the story regarding the quality of care in a particular setting or for a particular group of patients. Hard-copy medical records require several additional steps. First, sampled records must be obtained physically from where they are stored and, then, relevant data items from each record must be entered onto previously designed data-entry forms, and then the data from those forms must be extracted and entered on the computer for analysis. Each step creates another opportunity for errors to creep in. (Some of those steps can be combined. For example, extracted data can be entered directly onto a spreadsheet on a computer.)

4. Given These Limitations, What Can Reasonably Be Done With Utilization Data?

Even though utilization data, like all data have the limitations noted, they can be used to do at least five things of value: (a) determine whether or not services were

provided; (b) identify patterns of services provided; (c) determine whether follow-up services were provided when they were indicated; (d) include the values obtained from tests; and (e) include progress notes.

a. Utilization records can determine whether or not services were provided.

The core element in utilization records is an indicator that a service has been provided. Assuming there is evidence that a service was provided, the questions concern the level of detail incorporated in the record. Claims usually include simply that a physician visit occurred, a dilated eye exam was performed, or a prescription was written. They will not usually include an indicator that the patient's blood pressure was taken, what the doctor or nurse learned from the history, or the results of the eye exam. Thus, to use claims data in a study of quality is to be content with knowing the patient had a visit with his doctor or that a particular reimbursable test or other service was performed. If subsequent analysis suggests that, following treatment, patients with visits were healthier than patients with no visits or fewer of them (controlling for diagnosis and/or health status), then, further studies which investigate components of visits may be indicated. (More detailed studies could be done even if no differences in outcomes are observed to see if certain components of visits were beneficial even if the visit as a whole was not.)

Medical records, especially those that follow a prescribed format, can provide more detail. They can indicate not only what particular services were provided during a visit, but they can also include the nature of the tests that were done, the values on those tests (particularly useful if available before and after treatments), and details about the medications ordered. Although the medical record might not indicate whether the patient filled the prescription, claims will because, typically, the pharmacy is compensated when a prescription is actually filled. Neither will definitively record that the patient actually took the medication as directed unless the physician administered it himself (e.g., by injection) or observed the patient taking it (as in directly observed therapy often used in the treatment of tuberculosis).

Although with utilization data of this type, the delivery of a particular service can be associated with a specific measured outcome, usually that association is just assumed based on the medical literature. We cannot know what questions the physician asked, how sensitively he asked them, how well he heard what the patient actually said, or what the patient thought of the way in which the service was delivered. All of these are elements of quality, too, as defined above.

As HEDIS reports illustrate,[15] this kind of data can be used both to determine the extent to which certain specific services were provided to a plan's members as well as to examine the adequacy of care for members with certain diagnosed conditions. For example, it can calculate overall utilization measures like ambulatory visits per 1,000 member months or annual rates per 1,000 members. It can report the number of prescriptions per member per month or the average number per member per year. It can also report utilization measures for patients with particular conditions. These could include the proportion of patients with hypertension whose blood pressure was controlled or the proportion of patients who received beta blockers after heart attacks. All of these

results can also be adjusted for age or presented separately for different age groups as well as by gender or other demographic characteristics that are substantively useful.[16] Plan officials can compare their results to national averages or to those of competitor plans as part of their effort to assess the state of their own plan's performance.

b. Can identify patterns of services provided.

In addition to simply recording that a service or a series of services was provided, it is also possible to use utilization data to construct *patterns of care*. A visit to a primary care physician (PCP) is usually thought of as a good thing for patients with any number of chronic or acute conditions. The proportion of patients with those conditions who see a PCP in a year is often reported as a measure of quality on the presumption that during that visit, the physician was able to monitor the patient's progress and, if necessary, adjust the treatment regimen. On the other hand, while assumptions can be made about specific services that occurred during those visits and what other services were provided as a result of them, knowing only that a visit occurred omits a lot.

It is possible to go further, however, based on two assumptions: First, that for each diagnosis, an optimal pattern of care for most patients exists. It might not fit all such patients because of differences in severity, individual responses to treatment, or other factors; but other things being equal, a pattern of care can be described that is appropriate for most patients. The second assumption is that patients may obtain some but not all of the recommended services and, therefore, it is reasonable to think of gradations of quality instead of simply the presence or absence of a set of appropriate services.

In an earlier study, we created five quality categories for each of three chronic conditions based on the use of specific services appropriate to those conditions. The rationale for specific categories can be illustrated by the following example. The five components in the measure of good quality care for patients with diabetes are:

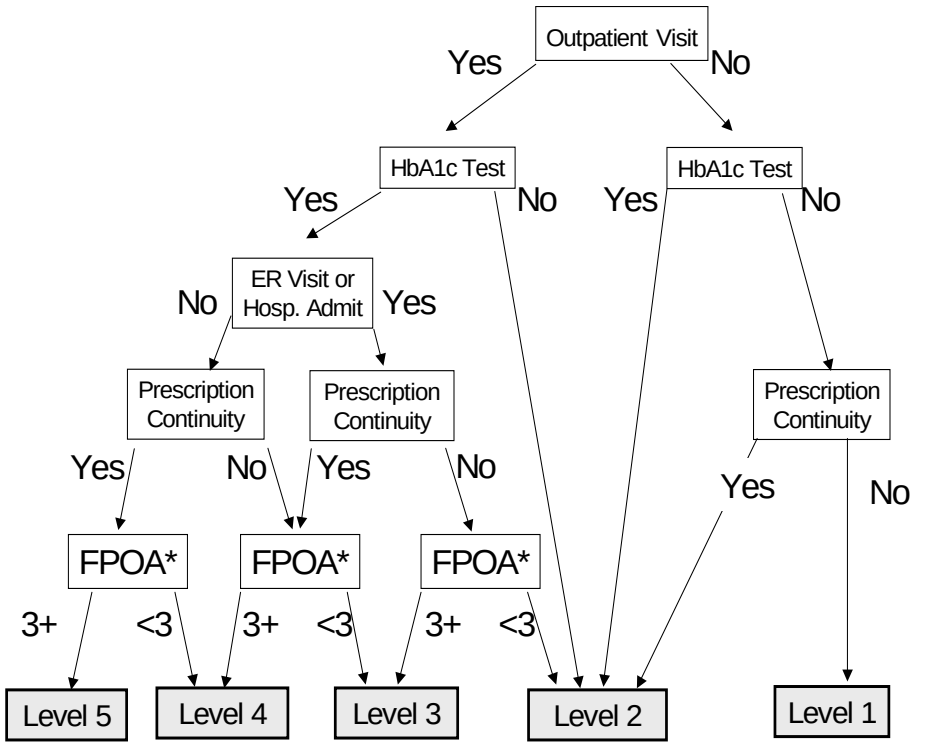
1) Having an outpatient visit during the year.
2) Receiving a hemoglobin A1c (HbA1c) test during the year.
3) Having continuity in hypoglycemic medications for more than 6 months.
4) Avoiding an ER visit (if not soon followed by an outpatient visit) or a diabetes-related hospitalization.
5) Receiving three or more of the following: (a) flu shot, (b) podiatry visit, (c) ophthalmology visit, or (d) microalbuminuria test or treatment with Angiotensin-converting Enzyme Inhibitors (ACEI) or Angiotension II Receptor Blockers (ARB).

With these services, we created five categories of quality and called them *level one care* through *level five care*, with the higher levels being preferred to the lower levels. (See Figure One for an example of the decision algorithm. The Appendix Table has detailed definitions for each category by diagnosis.) The categories or levels for patients with diabetes are as follows:

- *Level one* diabetes care means having no outpatient visit, no HbA1c test, and no continuity of hypoglycemic medications.

- *Level two care* means having only one of an outpatient visit, an HbA1c test, or medication continuity, but nothing else. While patients in this category are receiving some services aimed at their diabetes and, therefore, have somewhat better care than those in *level one*, it is still suboptimal because their condition is not being monitored adequately.

- Diabetes patients in the *level three care* category did receive the primary services of outpatient visit and HbA1c test. However, they either lacked other services (medication continuity or ancillary services) or had a diabetes-related ER visit without follow-up outpatient visit or a diabetes-related hospitalization. Care remains less than optimal because so many recommended services were not used.



*F=Flu Shot; P=Podiatry Visit; O=Ophthalmology Visit; A = microalbumin test or ARB or ACE inhibitor

Figure: Decision Algorithm for Assigning Quality of Care Level to Patients with Diabetes

- Patients with *level four care* received most of the recommended services, but lacked one element of the full complement of services that would constitute optimal care.

- *Level five care* means having used all the desired services and having had no ER visit or hospitalization for a diabetes-related reason that would show inadequate management of the condition.

Patients with a chronic condition should see a primary care physician (PCP) at least once a year for two main reasons: (1) to monitor their condition in order to avoid deterioration or retard its advance into something much worse; and (2) because visits to PCPs are the preferred point of entry into the health care system so that the PCPs, in addition to providing care directly, can guide their patients to services that are necessary and avoid those that are not.[17] For these and other reasons, managed care organizations (MCOs) often give subscribers relatively easy access to their PCPs who, in the case of people with a chronic condition, may be the specialist who treats it.

On this rationale, we gave greatest priority in diabetes management to outpatient visits and HbA1c tests. Those patients lacking an outpatient visit, HbA1c test and medication continuity were defined as *level one*, while patients with at least one of these were *level two*. The care of patients with an outpatient visit and an HbA1c test who met all of the other quality components (medication continuity, avoidance of an ER visit or hospital admission, and ancillary services) was defined as *level five* care. Those patients with an outpatient visit and an HbA1c test were moved down one category from *level five* for each one of these three other quality components that they lacked, including those patients who were defined as *level two* because all they had was an outpatient visit and an HbA1c test (i.e., three steps down from *level five*). These criteria are quite similar to those developed by the Diabetes Quality Improvement Project (DQIP) of the National Committee for Quality Assurance (NCQA)[18] and by HealthPartners of Minneapolis.[19]

Measures like these based on utilization data have several benefits, including the following: first, it is more reflective of an appropriate multi-step *process* of care that clinicians can relate to rather than simply the presence or absence of specific services; secondly, it permits leaders to identify particular deficits in that process to target for improvement; and finally, by doing so, it allows them to create intermediate, attainable goals and to measure progress toward the ultimate objective, not simply the attainment of a standard.

We can illustrate the potential value by referring to conclusions drawn from results in the study for which we created the levels-of-quality measures. First, large proportions of covered patients with all three chronic illnesses had utilization patterns that failed to meet condition-specific quality criteria. This finding, which is consistent with those reported by others,[20 21] demonstrates the need for improvements. It also shows the problem pervades the system, and was not caused by a few ineffective individual physicians. Secondly, the patterns tended to persist from year to year. Patients who were in level one in one year were likely to be in the same category in the next year, as well. Similarly, most patients in the highest categories also were likely to be in the same category in the next year. This finding shows that, whatever quality-improvement strategies they were using, leaders of the health plans from which the data were drawn were not having a positive impact.

c. Can determine whether follow-up services were provided when they were indicated.

One measure of quality surely is whether patients whose clinical condition warrants it receive follow-up care. Are patients who had operations seen by the surgeon after their surgery to determine whether the underlying problem was fixed, whether related new ones surfaced, or whether post-operative pain has subsided? Are patients who were treated for an acute illness better following the expected course of medication? Patients who no longer have the symptoms that led them to seek care may believe they no longer need service, and therefore, may not appear even if follow-up appointments were scheduled. Reasons include the cost in money, time, and other resources. In many cases, the patient may be right. He or she is better, a follow-up visit is unnecessary, and that is not an appropriate standard of care. Nonetheless, for patients with some conditions, physicians may want to schedule a visit to ensure that they have progressed as expected even though there is a very low probability that they have not. The challenge is to know for which patients follow up is appropriate and for which it is unnecessary. If such a visit is appropriate, then, it will be important to understand why it was not used. This is the case because the several parties have different interests, at least in the short run. The doctor tends to earn a fee for a follow-up visit and repeat tests; the insurer pays when those services are provided, but need not do so when they are not provided; and the patient wants to be better, but a follow-up visit has costs for him as well, whether or not the insurer pays some or all of the charge. Well designed studies using utilization data can determine whether follow up visits have value with particular types of conditions.

d. Can include details from patient visits.

Medical records may include details from patient visits that are likely to be missing when claims are the primary data source. These details can be in the form of values from the tests performed and progress notes from doctors and nurses entered during or shortly after a visit. When the values from tests are included, doctors and analysts can determine with greater precision the extent to which progress has occurred in the patient's condition. Without such data from medical records, analysts must depend on subjective patient reports – through surveys or other means – which require patients to remember what they felt like before the surgery and to articulate how they feel afterward. Test values are objective and can be collected during a visit without biasing the results. (A survey question asking a patient how he feels may generate a different answer when asked during a visit than a week later when the question asks him to remember what he felt like then.)

Contemporaneous progress notes entered by doctors and nurses about what the patient said, of their own impressions of the patient's condition, and of the possible diagnoses considered and treatments attempted can be especially valuable. Because such notes show evidence of a patient's condition at a particular point, subsequent entries can show change. While many of these notes will be subjective, they can be useful in assessing an individual patient's progress. In addition, when recorded over time and included with similar notes by other clinicians about other patients with the same condition, they can be very helpful in assessing the value of particular services.

5. What caveats must a careful analyst be aware of?

The value of data drawn from both claims and medical records is clear from this discussion. An important problem is that, while claims typically are available electronically and, therefore, retrievable with relative ease, medical records tend to be on paper and, thus, require much more effort to obtain. Once acquired, moreover, then, the data need to be entered into a computer file so that they can be analyzed, which helps to account for why it is done so infrequently. The spread of electronic medical records will make it more feasible to combine data from claims and medical records in the future.

Although utilization data from either source can be very useful in assessing quality of care and as a tool for making improvements, the discussion to this point also demonstrates the need for caution in using them. These are important caveats to be considered when using utilization data:

- Understand the data's limitations – including the potential for systematic inaccuracies resulting from their primary uses.

- Use care in transcribing to an analysis file – The quality of the source data is predetermined and cannot be changed, but entering it accurately for use in the study of quality is critical. Computer algorithms can be created to test for inconsistencies which will remove the most egregious errors.

- Understand the nature of the sample – Know how the groups are defined; how large and representative the samples are and the statistical power; and what conclusions are reasonable and what are not.

- Do not over interpret the results.

PART TWO: WAYS OF USING UTILIZATION MEASURES TO IMPROVE QUALITY AND SAFETY[22]

To change the way care is delivered in order to improve quality and safety requires the active involvement of clinicians, especially physicians. They are the clinical experts, and they are responsible for delivering services. If asked, most would say they are doing the best job they can and would be surprised, if not offended, at the suggestion that they could provide care in a better way. So, even if they can be persuaded that the organization faces an important problem, they are likely to believe that they cannot do anything to improve the situation. Therefore, getting them to the point that they will invest of themselves in improving service delivery will require several steps.

First, managers will need to get clinicians to believe three things: one is that whatever the level of quality and safety in the organization, they can be improved; the second is that the consequences of not dealing with it will affect not only the

organization, but also themselves; and the third is that, even though they may not be to blame, what they do as individuals can help to improve things. The best, if not the only way to get to that point is with data using utilization measures that show either that (1) summary results for your organization do not measure up to those of other organizations or to evidence-based standards, or (2) there is a lot of variation among individuals and/or groups within the organization, which means that while some may be practicing the highest quality medicine, others must not be.

Many doctors think of themselves as scientists and are influenced by data. In addition, like everyone else, they are influenced by incentives. So, if you have data generated by good studies that show care is not up to par; can show that, since the problem permeates the organization, the fault must lie with the system; and either can demonstrate that they are at risk, especially of losing income, or can appeal persuasively to their professionalism and commitment to patients, you have an opportunity to improve care.

Doing those things will only succeed in getting their attention, however. But with the data, you can begin a process to engage them in defining specific challenges and finding solutions to them.

6. Obstacles

Managers wanting to change the processes of care face a number of obstacles:

1. Differences in orientation of managers and doctors – Managers have responsibility for the organization; doctors focus on serving their patients one at a time.

2. Differences in connection to the organization – Managers are employed by the organization to lead it to success; most doctors are not employees but deliver services under the organization's auspices and may have similar connections with other organizations. They may have patients covered by several insurers or managed care organizations, and they may admit patients to several hospitals.

3. Partly as a result of the first two factors, doctors tend to distrust managers whom they see as trying to place constraints on their autonomy.

4. At the same time, managers are often reluctant to engage physicians on clinical issues because they are not their expertise.

5. Doctors tend to consider non-clinical activity, like meetings about administrative issues, to be a waste of time and a diversion from seeing patients.

7. Positive Factors

Despite these obstacles, the fact remains that efforts to change patterns of services require the involvement of doctors, so managers must find ways to engage them in that attempt. Fortunately, they have some assets on which to draw.

1. Information is a major asset because it can help you keep the conversation from being a conflict between management demands and clinical imperatives. You can frame the issue as a problem with the system and that, since they are the clinical experts, they are the ones who know enough to be able to solve it.

2. The payoff from success will be patients who will have better results and, therefore, be more appreciative of the clinicians; processes of care that will be smoother and more reliable; an organization that will be on sounder footing; and perhaps, greater financial rewards (or at least fewer costs).

3. An intangible benefit of success will be everyone's sense of accomplishment from having solved a difficult challenge as a team. This may be a hard sell going in, but all participants are likely to appreciate it if success is achieved. It should make it easier to get participation in the next problem-solving effort.

4. The "pain" of the process for participants, especially for non-salaried physicians, can be eased if it is possible to either compensate them or hold them harmless financially for their participation (if not in dollars for hours put in, then, perhaps in some creative way).

8. The Process

Theory shows that specific recommendations need to be geared to the particular organization and situation. Formulas do not work. Specific change efforts must be built on careful analysis and on diagnosis of conditions in each organization.

Step One, Organizational Assessment – The change effort can begin with an analysis of utilization data that will help you define the problem. Start with several diagnoses and determine the extent to which care provided in the organization meets evidence-based standards. In addition, pinpoint as best as possible those elements of the care process that are the most problematic. Data from several diagnosis that show a similar pattern may provide important clues. If possible, advance some hypotheses or tentative suggestions as to changes that can be made.

The Organizational Assessment should also include an understanding of the state of relations between clinicians and management as well as at least a beginning understanding of the processes of delivering care to patients with the focal diagnoses that develops from a systematic process analysis of the elements of care for those patients. More detailed understanding of those processes and strategic change points can develop out of the participatory process involving doctors and other clinicians. They really are

the experts, and they know both where their own frustrations lie and what they need to make clinical practice more rewarding.

Step Two, Picking a Place to Start – From among the problems identified by the utilization analysis, choose a specific focus (or two or three) for action. These efforts take time and require the efforts of many people, so keep them manageable. Define the processes of care in detail. Identify hypotheses to explain why too many patients do not get services indicated by their condition or why an episode of care takes longer than it needs to take. Remember that part of the reason may be that patients do not come in for services when they should and that some of the reasons for that result might be in the incentives imbedded in their insurance coverage (e.g., copayments or coinsurance) or in other costs like the lost income from taking time off from work. Part of the analysis should also include the extent to which perspectives on conditions vary by job title, position in the organization, or profession. Patients' perspectives can contribute useful insights, as well.

Step Three, Choosing a Direction for Change – Following the analysis in step two, identify particular targets for change and develop detailed action plans with the involvement of key stakeholders. Guidelines for leading change can include the following:

1. Put together a multi-disciplinary group to lead the effort.

2. State the objectives of the change initiative clearly and let the affected parties comment. Be willing to modify goals. Do not overpromise.

3. Share information. Make results of initial assessments available to all relevant parties; listen to their reactions; and respond to them. People will be more likely to support change if they develop a sense of shared ownership in it that comes from providing input that is valued by others.

4. Don't work uphill. Pick projects that are real (don't waste people's time), and that are likely to succeed. Use leaders who are respected by large numbers of participants. Charge the group members with working through the problems. Provide leaders with training and/or consultation in group problem solving.

5. Set a timetable for achieving milestones so that expectations are clear and the group feels some pressure to achieve results. On the other hand, if it becomes necessary to renegotiate it, do so publicly, out in the open.

CONCLUSION

Quality of care is defined by the outcomes it produces. Utilization data are, therefore, essential tools in, first, assessing the state of quality in a given health care organization and, then, in developing a strategy to improve it. Then, the conclusions derived from the analysis of those data can be used to begin a process of change that must

involve the people whose activities must change in order for the organization to show improved results. There is no way to avoid the change process. It takes time and it can be messy, but it will rest on a strong foundation if the organization's leaders understand the message in their utilization data and continue to use it throughout that process to set goals and measure accomplishments.

APPENDIX

Quality Categories for Patients with Asthma, Diabetes, and Congestive Heart Failure

Disease and Quality Categories	Quality Criteria
ASTHMA*	
Level one care:	• No PCP visit, no inhaled steroid
Level two care:	• No PCP visit, but inhaled steroid
Level three care:	• PCP visit, no inhaled steroid, ER visit
Level four care:	• PCP visit, inhaled steroid, ER visit or • PCP visit, no inhaled steroid, ER visit
Level five care:	• PCP visit, inhaled steroid, no ER visit
CONGESTIVE HEART FAILURE**	
Level one care:	• No PCP visit, no ACE Inhibitor continuity
Level two care:	• No PCP visit, no ACE Inhibitor continuity or • PCP visit, ER visit, no ACE Inhibitor continuity, and less than 2 of the following: flu shot, beta blocker, and assessment of left ventricular function
Level three care:	• PCP visit, ER visit, no ACE Inhibitor continuity, and 2 or more of the following: flu shot, beta blocker, and assessment of left ventricular function • PCP visit, ER visit, ACE Inhibitor continuity, and fewer than 2 of the following: flu shot, beta blocker, and assessment of left ventricular function • PCP visit, no ER visit, no ACE Inhibitor continuity, and fewer than 2 of the following: flu shot, beta blocker, and assessment of left ventricular function
Level four care:	• PCP visit, ER visit, ACE Inhibitor continuity, and 2 or more of the following: flu shot, beta blocker, and assessment of left ventricular function • PCP visit, no ER visit, no ACE Inhibitor continuity, and 2 or more of the following: flu shot, beta blocker, and assessment of left ventricular function • PCP visit, no ER visit, ACE Inhibitor continuity, and fewer than 2 of the following: flu shot, beta blocker, and assessment of left ventricular function
Level five care:	• PCP visit, no ER visit, ACE Inhibitor continuity, and 2 or more of the following: flu shot, beta blocker, and assessment of left ventricular function
DIABETES***	
Level one care:	• No PCP visit, no HbA1c test or • No PCP visit, no HbA1c = 0, no continuity of hypoglycemic medications
Level two care:	• No PCP visit, no HbA1c test, continuity of hypoglycemic medications or • PCP visit, no HbA1c test or • No PCP visit, HbA1c test or • PCP visit, HbA1c test, ER visit, no continuity of hypoglycemic medications, and 2 or fewer of the following 4: flu shot, podiatry visit, ophthalmology visit, and ACE/ARB inhibitor or microalbuminuria
Level three care:	• PCP visit, HbA1c test, no ER visit, no continuity, and 2 or fewer of the following: flu shot, podiatry visit, ophthalmology visit, and ACE/ARB inhibitor or microalbuminuria • PCP visit, HbA1c test, ER visit, no continuity of hypoglycemic medications, and 3 or

Level four care:	more of the following: flu shot, podiatry visit, ophthalmology visit, and ACE/ARB inhibitor or micro-albuminuria • PCP visit, HbA1c test, ER visit, continuity of hypoglycemic medications, and 3 or more of the following: flu shot, podiatry visit, ophthalmology visit, and ACE/ARB inhibitor or microalbuminuria • PCP visit, HbA1c test, no ER visit, continuity of hypoglycemic medications, and 2 or fewer of the following: flu shot, podiatry visit, ophthalmology visit, and ACE/ARB inhibitor or microalbuminuria • PCP visit, HbA1c test, no ER visit, no continuity of hypoglycemic medications, and 3 or more of the following: flu shot, podiatry visit, ophthalmology visit, and ACE/ARB inhibitor or microalbuminuria • PCP visit, HbA1c test, ER visit, continuity of hypoglycemic medications, and 3 or more of the following: flu shot, podiatry visit, ophthalmology visit, and ACE/ARB inhibitor or microalbuminuria
Level five care:	• PCP visit, HbA1c test, no ER visit, continuity of hypoglycemic medications, and 3 or more of the following: flu shot, podiatry visit, ophthalmology visit, and ACE/ARB inhibitor or microalbuminuria

***DEFINITIONS FOR ASTHMA:**

ER = either (1) ER visit for asthma or (2) hospitalization for asthma.

PCP: CPT4 codes 99201-99205, 99211-99215, 99241-99245 and outpatient visit with a PCP or pulmonologist or allergist.

****DEFINITIONS FOR CHF:**

ER = either (1) ER visit and NO outpatient visit within a month or (2) hospitalization for CHF-related condition . CHF-related conditions are CHF, anasarca, or cardiomyopathy.
Continuity: ACE Inhibitor for at least 6 months.

PCP: CPT4 codes 99201-99205, 99211-99215, or 99241-99245 and outpatient visit with PCP or cardiologist.

F = Flu shot
B = Beta blocker
L = Assessment of left ventricular function. Defined as either of these CPT-4 codes for laboratory tests: 93320 (echocardiogram) or 01920 (cardiac ventriculogram).

*****DEFINITIONS FOR DIABETES:**

ER = either (1) ANY ER visit and NO outpatient visit within a month or (2) hospitalization for diabetes-related condition. Diabetes-related conditions are foot ulcer, cellulitis, hypoglycemia, or diabetic ketoacidosis.

Continuity: hypoglycemic medications for at least 6 months.

PCP = CPT4 codes: 99201-99205, 99211-99215, 99241-99245 and PCP or endocrinologist.

F = Flu shot
P = Podiatry visit
O = Ophthalmology visit
A = ACE/ARB inhibitor or microalbuminuria

[1] Leatherman, S. and D. McCarthy. Quality of Health Care in the United States: A Chartbook. New York: The Commonwealth Fund, April 2002.
[2] Ibid., p. 25.
[3] Ibid. p. 36.
[4] Ibid. p. 40.
[5] Ibid. p. 58.
[6] National Committee for Quality Assurance. *The State of Health Care Quality 2004: Industry Trends and Analysis.* Washington, D.C.: NCQA, 2004., p. 9.
[7] Loc. cit.
[8] D. M. Eddy, 1998. Performance Measurement: Problems and Solutions. *Health Affairs* 17(4):7-25 at p. 8.
[9] E. A. McGlynn, 1997. Six Challenges In Measuring The Quality of Health Care. *Health Affairs* 16(3):7-21.
[10] D. M. Eddy, 1998. Performance Measurement: Problems and Solutions. *Health Affairs* 17(4):7-25 at p. 8.
[11] E. A. McGlynn, 1997. Six Challenges In Measuring The Quality of Health Care. *Health Affairs* 16(3):7-21.
[12] D. Blumenthal, 1996. Quality of Care: What Is It? *NEJM*, (Sept. 19):891-894.
[13] K. N. Lohr, ed. *Medicare: A Strategy for Quality Assurance.* Washington, D.C.: National Academy Press, 1990. Quoted in M. A. Schuster, E. A. McGlynn, C. B. Pham, M. D. Spar, and R. H. Brook. The Quality of Health Care in the United States: A Review of Articles Since 1987. in Institute of Medicine *Crossing the Quality Chasm: A New Health Ssytem for the 21st Century,* Washington, D.C.: National Academy Press, 2001.
[14] A. Donabedian……
[15] National Committee for Quality Assurance. *The State of Health Care Quality 2004: Industry Trends and Analysis.* Washington, D.C.: NCQA, 2004.
[16] For an example, see reports available on the website of the Minnesota Department of Health, which presents such reports separately for HMOs doing business in the state. http://www.health.state.mn.us/divs/hpsc/mcs/hedis.
[17] B. Starfield, Primary Care: Balancing Health Needs, Services and Technology, Second Edition (Oxford: Oxford University Press, 1998).
[18] NCQA/DQIP website at www.ncqa.org/dprp/dqip2.htm.
[19] HealthPartners. Clinical Quality Measures: Optimal Care for Diabetes. Found at http://www.healthpartners.com. Click on Compare and Select Providers and then on View Quality of Care and Satisfaction Comparisons.
[20] HealthPartners, Clinical Quality Measures: Optimal Care for Diabetes. Found at http://www.healthpartners.com. Click on Compare and Select Providers and then on View Quality of Care and Satisfaction Comparisons.
[21] E.A. McGlynn, S.M.Asch, J. Adams, J. Keesey, J. Hicks, A. DeCristofaro, E.A.Kerr, The Quality of Care Delivered to Adults in the United States. *NEJM*, v. 348, no. 26, June 26, 2003, pp. 2635-2645.
[22] This section is based on S. M. Davidson, M. McCollom, and J. Heineke, *Building a Health Care Organization: A Challenge for Physicians and Managers.* Washington, D.C.: Beard Books, 2005. (Originally published as *The Physician-Manager Alliance: Building the Healthy Health Care Organization.* San Francisco: Jossey-Bass, Inc., 1996.)