Quality Improvement Initiatives

Issues in moving from diabetes guidelines to policy

RODNEY A. HAYWARD, MD1,2,3
TIMOTHY P. HOFER, MD, MSC1,2,3
EVE A. KERR, MD, MPH1,2
SARAH L. KREIN, PHD, RN1,2

OBJECTIVE — To outline the principles that direct the Veterans Affairs (VA) Quality Enhancement Research Initiative (QUERI) dedicated to diabetes quality improvement (QUERI-DM).

RESEARCH DESIGN AND METHODS — We discuss the VA initiatives aimed at improving diabetes care for veterans as well as general issues that should be considered in quality improvement initiatives. We specifically describe some of the epidemiological, statistical, and organizational issues that have guided our quality improvement (QI) programs.

RESULTS — The five principles that have guided the QUERI-DM process are: 1) treating clinical guidelines and goals distinct from quality standards and quality improvement priorities; 2) targeting high-risk patients and high-impact quality issues; 3) profiling processes over outcomes; 4) targeting processes that will improve patient outcomes; and 5) paying attention to the loci of practice variation.

CONCLUSIONS — The authors recommend that all five principles be considered when moving from practice guidelines to performance measures and QI initiatives. Targeting high-priority problems and high-risk groups can greatly improve the effectiveness and efficiency of QI interventions.

Diabetes Care 27 (Suppl. 2):B54–B60, 2004

The obvious, and most important, policy issue in selecting a standard of quality is the social commitment to implement that standard, not as a means for passing judgments which can then be ignored, but as a basis for making available whatever is needed to make the standard an achievable objective. Failing that, the standard is either a phantasm that can be safely ignored, or an intrusive presence that must be exercised.—Avedis Donabedian (1).

The large gap between the care that many people with diabetes receive and optimal practice has been well chronicled (2–6). Deficiencies in care result in preventable complications, premature mortality, and many millions of wasted health care dollars annually (7–10). Yet, we know relatively little about the reasons for suboptimal care and even less about the best policies for enhancing care. In this article we discuss some of the Veterans Administration (VA) initiatives aimed at improving diabetes care for veterans, as well as several broader issues that should be considered in quality improvement initiatives. In particular, we discuss some of the epidemiological, statistical, and organizational issues that have guided our quality improvement (QI) programs. We propose that these issues should be considered when transitioning from practice guidelines to performance measures, quality improvement initiatives, and other policy interventions.

In recent years, the VA has been extremely active in promoting increased provider accountability and initiating QI initiatives. For example, diabetes quality measures are included in ongoing chart audits conducted for the VA Office of Quality and Performance (11), and the VA Healthcare Analysis and Information Group, formerly the National Center for Cost Containment, has been collecting and reporting data on pharmaceutical use and glycemic control since FY1994 (12). In addition, VA Headquarters’ Office of Patient Care Services oversees a variety of initiatives (including a national amputation prevention program) (VA Directive 96-007) and has led a diabetes care guideline development and dissemination initiative (13). Many of these programs are highlighted elsewhere in this special issue of Diabetes Care.

In 1998–1999, the VA commissioned the Quality Enhancement Research Initiative (QUERI) to help facilitate QI in VA through eight disease-specific QUERI initiatives (14), one of which is dedicated to diabetes QI (QUERI-DM) (15). QUERI-DM is charged with bringing additional epidemiological, economic, and statistical expertise to VA QI programs by increasing interactions among researchers, clinicians, management, and others involved in quality assessment and improvement (15). QUERI-DM is also charged with investigating important gaps in our knowledge, as well as conducting rigorous evaluations of the effectiveness of ongoing and new QI initiatives.

The QUERI process follows a step-by-step approach to QI (see Table 1). Each
Table 1—QUERI-DM objectives and some examples of initiatives

<table>
<thead>
<tr>
<th>QUERI steps</th>
<th>Examples of QUERI projects</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Identify best practices</td>
<td>Cost-effectiveness analysis of different intervals for retinal screening and for use of new diabetes medications.</td>
</tr>
<tr>
<td>2. Define existing practice patterns and outcomes</td>
<td>Compile a coordinated diabetes database and reporting system. Validate the reliability of quality profiles using computerized medical information systems. Examine the most common causes of suboptimal timing of retinal photocoagulation.</td>
</tr>
<tr>
<td>3. Identify and implement interventions to improve practice</td>
<td>Evaluate the effectiveness of a beeper reminder system in improving self-care for hyperglycemia. Physician and clinic director feedback and profiles on patients with persistently poor HbA1c and LDL values. Examine whether adoption of specialized foot clinics results in fewer amputations.</td>
</tr>
<tr>
<td>4. Document that best practices improve outcomes</td>
<td>Examine whether adoption of interventions to change provider behavior, development of new systems to make providing care easier and more fail-safe, and issues associated with guideline development and implementation. One of the key steps in QI is selecting the key aspects of care to focus on. In this article, we discuss five principles (outlined in Table 2) that have helped guide us in the QUERI-DM process. These issues are critical in selecting important QI projects, in deciding what measures to monitor and report on, and in evaluating whether a QI intervention has worked.</td>
</tr>
</tbody>
</table>
| 5. Document that outcomes are associated with improved health-related quality of life | Pay careful attention to the loci of practice variation. Adoption of best practices can be slow. Past experience suggests that targeted implementation strategies are necessary to obtain substantial improvements in guideline adherence (17–20). If promoting cost-effective QI is the main policy goal, treatment goals set forth in guidelines should rarely be adopted, without modification, as accountability measures, quality standard profiles, or provider profiles (i.e., “report cards”) (1). It is essential that we keep in mind that the definition and purpose of guidelines and those of quality standards are different. While guidelines are intended to be comprehensive, flexible, and instructive, quality standards should measure past clinical actions that have a strong process-outcome link (21). Well-developed guidelines result from assessing the available scientific evidence and arriving at usual recommendations and treatment goals for care. In an effort to be specific and comprehensive, guidelines often include recommendations for which there is limited evidence or for which the intervention will have a small impact on outcomes. Cost-effectiveness is often not considered because the treatment goal is an ideal that is often not expressed in terms of one specific treatment. For example, the guideline may list a usual HbA1c goal of <7%; however, the guideline’s authors may fully understand that the desirability of changing therapy in pursuit of that goal may depend on what treatments are required to achieve the goal. Individual patient circumstances and sound clinical judgment may dictate that deviation from the guidelines is the most appropriate medical care. As the term implies, “guidelines” are simply tools that should be used intelligently to help set individualized goals by providers and their patients and should not be considered either a maximum or a minimum level of care. Quality standards should represent high priorities for medical care for which an institutional or societal commitment to produce this level of care exists. While these standards may not measure all as-

Table 2—Five key factors in choosing and implementing QI initiatives

1. Do not confuse guidelines and goals with quality standards.  
2. Target high-risk patients and high-impact quality issues.  
3. Profiling processes is usually preferable to profiling outcomes.  
4. Only target processes that are very likely to improve patient outcomes.  
5. Pay careful attention to the loci of practice variation.
Table 3—QUERI-DM’s highest priorities for diabetes care improvement

<table>
<thead>
<tr>
<th>Priorities for the care of the average type 2 diabetic patient</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>● HbA1c &lt;9.0–9.3%* (23)</td>
<td>● LDL cholesterol &lt;130–140 mg/dl† (41, 42)</td>
</tr>
<tr>
<td>● Blood pressure &lt;135–140/80–85 mmHg† (42–44)</td>
<td>● Aggressive foot care for high-risk patients* (45, 46)</td>
</tr>
<tr>
<td>● Optimally timed photocoagulation for diabetic retinopathy† (7, 10, 22, 47)</td>
<td>● Daily aspirin use† (42, 48, 49)</td>
</tr>
<tr>
<td>● Smoking cessation (50)</td>
<td>● Promoting at least moderate physical activity† (51, 52)</td>
</tr>
<tr>
<td>● A heart healthy diet† (53)</td>
<td></td>
</tr>
</tbody>
</table>

Additional priorities for type 1 and early-onset type 2 diabetes care

<table>
<thead>
<tr>
<th>Priority for type 1 diabetes care</th>
<th>Priority for type 2 diabetes care</th>
</tr>
</thead>
<tbody>
<tr>
<td>● HbA1c &lt;8.0%* (23, 54)</td>
<td>● Use of ACE inhibitors to slow progression of diabetic nephropathy* (55, 56)</td>
</tr>
</tbody>
</table>

The best available epidemiological evidence strongly suggests that each of the above will have a substantial impact on end-stage patient outcomes (such as vision loss, stroke, or premature death), however, the nature of the evidence varies: *Randomized trials have established a causal link with intermediate outcomes (such as disease progression), and statistical modeling and extrapolation based on combined clinical trial and observational studies suggest that a major improvement in patient outcomes would be expected with longer follow-up. †Randomized trials have established that the intervention has a major impact on end-stage outcomes in high-risk patients, but there has not yet been definitive diabetes-specific subanalyses. ‡Randomized trials have established that the intervention has a major impact on end-stage outcomes in patients with diabetes. §Very strong epidemiological evidence, but no clinical trial has been conducted.

Aspects of a guideline implementation intervention, they will assess whether the intervention influenced the most important processes of care. Poorly conceived accountability measures can waste precious resources by diverting effort and money that could have been used on other, potentially more worthy, QI initiatives. Both the direct costs of data collection and analysis and possible incentives for providing discretionary care (such as requiring annual eye exams in a diabetic patient whose last examination was normal [22]) can be very wasteful. However, the costs of QI are not limited to the time and money of the intervention and care itself, but also include the consumption of other scarce resources, such as clinicians’ and patients’ time and attention. The attempt to improve 10 aspects of care may be less effective for each of the 10 compared with targeting a single item for QI. Consequently, each intervention targeted should be important enough not only to justify the direct expense of that initiative, but also worth the risk of diluting the impact of concurrent QI efforts and diverting funds from higher-priority quality problems.

While clinicians may use guidelines to direct their daily practice and make patient recommendations, it is rarely appropriate for administrators to use guidelines directly as accountability or quality measures to monitor past performance. In identifying how to develop quality standards to assess the success of a guideline implementation project, the following should be considered: 1) the weight and strength of the evidence that improving an aspect of care will impact important patient outcomes (some parts of guidelines are based on stronger evidence than others), 2) the absolute risk reduction that will be achieved if care is improved to meet the standard, 3) a clinician’s likely response to implementation of the measure (could the accountability measure lead to perverse incentives to game the system by patient deselection or referral), 4) whether there are distinct subpopulations (usually defined by presence of risk factors) that may require less or more intensive treatment, and 5) whether most patients would want the recommended medical intervention if they were well informed.

2) Target high-risk patients and high-impact quality issues

In selecting QUERI-DM initiatives we have tried to concentrate on the most critical quality problems (Table 3). Even if we all agree that maintaining an HbA1c <7% is a good goal, having an HbA1c of 8% is a small problem compared with having one of 10% (23). Certainly, initiating more intensive and expensive pharmacologic treatment in some patients with an HbA1c of 8% may not be cost-effective and may also not be desired by the patient after considering the added monetary costs, increased glucose monitoring, and potential side effects. Recognizing that resources for QI are scarce, and that pressures for cost containment are high, it may be more appropriate for a health care system to target patients with HbA1c values >9.0% for special attention and resources, at least for the average patient with type 2 diabetes (see Table 3) (24). It is those with poor control and early onset of disease who account for the majority of blindness and end-stage renal disease (23), and perhaps we should consider them the highest priority for scarce QI resources. Directing QI resources to high-risk patients should not be seen as encouraging providers to only reduce HbA1c values for their patients to <9.0%. First, current guidelines clearly state that tighter control is highly desirable and that continued educational efforts at communicating this to providers and patients are important. Second, clinicians still have a moral responsibility to inform their patients about the risks and benefits of treatment and work on optimal treatment plans with their patients. Third, given the variability in HbA1c measures (and many other accountability measures), leaving a patient close to the threshold increases the risk that the next measurement will not meet the performance standard.

Accountability measures should have substantial implications for poor performance, providing a strong incentive for providers to improve their performance. Setting accountability measures at tight levels (i.e., percent of patients with HbA1c <7.0% or LDL cholesterol <100 mg/dl) often results in debate about the appropriateness of the strict measure, and a great deal of effort is expended in arguing the value and feasibility of such tight control in all patients (i.e., trying to “exorcise” the measure). In addition, stringent quality and accountability standards could also have paradoxical adverse effects. At first glance, it may seem that an optimal performance standard would be to maximize the percentage of patients who have an HbA1c <7.0%. Such a standard may divert a provider or health system’s attention from treating poorly controlled patients to disproportionately focusing on the larger numbers of patients who are slightly above this cutoff. Indeed, this approach could encourage physicians to concentrate on those who have much lower baseline risk of end-stage complications (23) because they are already near
“tight control,” even though intensified treatment of these patients may be much less likely to substantially lower their HbA1c levels (6). If we set the accountability measure as the percent of patients <9.0%, we encourage providers to concentrate their efforts on those with the highest risk who will, on average, achieve the greatest improvement in HbA1c (although perhaps not “tight” control) with new treatments (6,23). Once a QI initiative has made substantial improvements for the highest risk group, the next priority might be to focus greater attention on promoting tighter control in patients with moderate HbA1c elevations or perhaps switch to an intervention on a different domain, such as blood pressure control.

3) Profiling processes are usually preferable to profiling outcomes

If QI should be directed at improving outcomes, why don’t we simply hold providers accountable for their patients’ outcomes by creating risk-adjusted outcome rates (i.e., amputation, blindness, hospitalization, or mortality rates)? While intuitively appealing, this approach is rarely appropriate. Hofer and Hayward (25) have demonstrated that even with superb case-mix adjustment, profiling individual hospitals based on their mortality rates is likely to be highly inaccurate, at least for medical admission diagnoses. In addition, primary care physicians can improve their diabetes profiles of patient glycemic control more easily by deselecting just one or two patients than by improving their care (26). This means that heavy-handed profiling of outcomes could encourage discrimination against the sickest and most needy patients. Certainly, risk-adjusted outcomes can be an important research or a preliminary exploratory tool; however, a policy of assessing quality based on hospital or provider-specific outcome rates actually runs counter to the implications of recent health services research (25–30).

First, it should not be surprising that outcome rates specific to hospital or providers often do not accurately reflect quality. Such rates may be based on small numbers that tend to be imprecise and variable (25–26). Furthermore, case-mix adjustment may not completely remove the influence of extraneous variables on outcome rates (27). This leads to policy debates and arguments about where differences in outcome rates are related to intrinsic quality differences as opposed to unmeasured patient severity. Even if perfect case-mix adjusters could be developed, which some authors believe is an unlikely event (26–29), collecting and analyzing the data to control for patient severity of illness, and monitoring for gaming or fraud, are both expensive and laborious.

However, once well-conducted clinical trials have shown that one treatment is substantively better than another, it is much easier to identify high- and low-quality providers by measuring provision of that treatment (a process) than by measuring the outcomes directly. In looking at care for acute myocardial infarction, Mant and Hicks (30) showed that it takes up to 10 times the sample size to distinguish hospitals based on outcomes (e.g., mortality rates) than that for evaluating the processes of care directly (i.e., receipt of thrombolytics, β-blockers, and ACE inhibitors in appropriate patients). Once we know that certain processes (e.g., aspirin once a day for diabetic individuals) or that intermediate outcomes (e.g., glycemic or blood pressure control) improve patient outcomes (e.g., visual impairment or mortality), then it is statistically more efficient to measure processes and intermediate outcomes (30,31). This is particularly important in diabetes when improvements or deficiencies in care may take many years, or even decades, to impact patient end-stage outcomes (e.g., visual impairment and kidney disease).

Recently, we have proposed using “tightly linked” quality measures in which the clinical intervention or process is strongly and directly linked to patient outcomes, an actionable process is measured, and a high-risk population is targeted (32). The eligibility for such a measure can be defined by the specific diagnosis or by a poor intermediate outcome (e.g., HbA1c level, blood pressure level, or LDL cholesterol level), but the measure itself is a process-of-care measure. For example, such a tightly linked measure could report the proportion of diabetic patients with a high LDL cholesterol level (e.g., >130 mg/dl) in the previous year and who, in the reporting year, were 1) started on cholesterol-lowering medication, 2) had the dose of their cholesterol-lowering medication substantially increased, or 3) had a repeat LDL level that was <130 mg/dl. Because tightly linked measures focus on the clinical intervention (e.g., giving appropriate treatment) rather than on the intermediate outcome only, providers get credit for appropriate care, regardless of the severity of their patient population. This can greatly mitigate or eliminate the incentive to avoid “sick” or nonadherent patients, is more likely to identify a high proportion of patients with genuine quality problems, and focuses attention on the QI intervention (process) known to improve care (32). In addition, our analyses thus far suggest that this type of measure exhibits much greater physician-level practice variation than that observed when examining LDL cholesterol or HbA1c levels in isolation, which will increase the reliability and meaningfulness of the profiles (33).

4) Only target processes that are very likely to substantially improve patient outcomes

Although process measures are often easier to track and are more statistically efficient for monitoring care quality, QI initiatives should always begin and end by considering important patient outcomes. It is sometimes too easy for us to believe in and advocate for the inherent virtues of our favorite interventions. In general, patient education, screening, early treatment, and the newest medications and technologies are only as important as the evidence that they actually improve patient outcomes. Too often processes have been monitored and referred to as “quality measures” mainly because they are easy to measure or seem like a good idea (32). Although well-designed clinical trials may be the best method for understanding which processes result in better patient outcomes, there is much that a coordinated health care system can do to better understand the potential relationship between its care processes and outcomes (31).

For example, consider early detection and treatment of diabetic retinopathy. The evidence certainly supports that early intervention with photocoagulation therapy will delay visual impairment (7,10,16,22). Screening for retinopathy detects this condition in its early stages and therefore facilitates appropriately timed treatment with photocoagulation therapy. Nevertheless, the optimal frequency of screening for this condition is unknown (22). Given this uncertainty, it may be reasonable to recommend, in clin-
ical guidelines, that annual screening is the safest approach, but the best available evidence suggests that annual screening (versus, for example, every other year) will not yield much benefit for the average patient. In the absence of an experimental trial to evaluate optimal screening intervals, how might an integrated health care system, such as the VA or a large HMO, determine whether an intervention to enhance screening, beyond that observed in current practice, should be considered. One approach, which QUERI-DM has initiated, is to examine the reasons that some patients do not get optimally timed photocoagulation (i.e., laser treatment at a point when delay in treatment has put the patient at risk of permanent visual impairment).

The current evidence clearly suggests that screening (efforts to detect the presence of early retinopathy) should be distinguished from surveillance (follow-up of those with known early retinopathy) (22,34). Because current policy focuses on screening, we might begin by evaluating the adequacy of our screening program by examining the circumstances of suboptimal timing of photocoagulation—not by merely determining how often we meet the screening interval recommended in the guidelines (which is based on weak evidence at best). For example, if we find that preventable visual loss frequently results from an interval between screening retinal examinations that is somewhat longer than 1 year (e.g., 2 years), then increasing routine annual screening frequency should be a high priority. If, on the other hand, presenting to ophthalmology with late-stage retinopathy is almost exclusively related to very long intervals between screening (≥4 years), but is unrelated to examinations 3 or fewer years apart, then the focus would shift to developing an intensive system that tries to ensure that no patients with diabetes go >3 years without screening. At this point, further investigation may be necessary to determine the reasons behind this 3- to 4-year gap between eye exams for this group (e.g., inability to get to the clinics during usual clinic business hours or travel distance to the clinics), so as to overcome important barriers. Given the latter scenario, an intensive intervention, targeting the 10–20% who have gone >3 years without an eye screening, would likely do more to decrease visual loss than a less intensive intervention directed at all patients.

However, our initial actual findings thus far (which need to be verified in larger, more representative samples) suggest that suboptimal timing of photocoagulation is mainly due to problems with surveillance for those with known early retinopathy (those who might benefit from 3- to 6-month between-exam intervals). If these pilot findings are confirmed by larger studies, then the predominant policy of concentrating on frequent screening (as opposed to surveillance and timely treatment of known disease) may do little to improve quality. Not only does the current policy ignore the issue of close surveillance, it could even exacerbate the problem by overcrowding ophthalmology outpatient services with mandated care that has little or no value. Whatever the final results find in this particular instance, we feel that when we mandate spending annually tens of millions of dollars on a “quality initiative” that we should also devote some resources to a systematic and rigorous evaluation of the patterns and circumstances of the preventable adverse events that we are trying to eliminate. Insights from such evaluations can be invaluable in successfully selecting and targeting appropriate QI initiatives.

5) Pay careful attention to the loci of practice variation
If a QI intervention is to be tailored to the reasons for suboptimal care, examining the characteristics (e.g., patient, doctor, group practice, clinic, hospital) associated with the most variation in care can help to inform what the intervention should be and to whom it should be directed. Often, great time and expense are exerted on generating provider profiles (“report cards”), even though little or no systematic variation in receipt of care exists between providers’ practices for the measure being profiled (26,35). When variation does exist between providers or hospitals, it may be due to patient case mix or severity but could also be related to systematic differences in medical practice, as well as organizational differences between sites. In many instances, discussions between providers can help determine whether certain differences are justifiable, and providers at both high- and low–event rate institutions can learn from talking about their approaches to care and their organizational structure and processes (36). QUERI-DM is currently investigating variations in important aspects of diabetes care (glycemic control and LDL cholesterol levels). Preliminary findings suggest that for most of these measures, there is much more variation between facilities than between individual providers at a given facility, and the expense of developing individual provider profiles is often not appropriate. However, in some instances, substantial provider variation is present and individual provider profiles can be reliably constructed (33).

Statistical techniques for assessing multilevel variation and adjusting profiles for chance variation have been well described by others (26,37–40). Hierarchical modeling techniques have become much more accessible, and no profiling endeavor should be undertaken without first establishing the maximum amount of variation attributable to providers (versus patients and sites) and the number of patients needed to create reliable profiles (26). If strong sanctions are applied to profiles at the physician level, when in fact there is little variation due to a physician effect, it creates an incentive for physicians to game the system in ways that may actually adversely affect patient care (26).

Resources for QI are scarce and precious. Although the goal of guidelines and provider education may be to aid providers in a quest for individualized optimal care, we feel that the role of QI should be to facilitate the greatest population-based improvements in patient outcomes with at least some consideration being given to economic efficiency. Ironically, improving outcomes may often be better accomplished by focusing on key processes rather than directly measuring risk-adjusted outcomes. In addition, QI will often be best accomplished by targeting high-risk subgroups (e.g., those with HbA1c >9.0%) rather than by excessively focusing on strict “guideline compliance,” which may result in little marginal benefit and may not be appropriate or achievable in many patients. We feel that working to understand the nature of a quality problem and then considering the five principles outlined above can help create much more effective and efficient QI initiatives as we all work together to prevent the devastating complications
of this important and treatable chronic condition.

Acknowledgments — This work was supported by the VA Quality Enhancement Research Initiative (QUERI). E.A.K. is supported by a VA Health Services Research & Development Career Development Award. Additional support was provided by the Michigan Diabetes Research and Training Center Grant from the National Institute of Diabetes and Digestive and Kidney Diseases (P60DK-20572).

References